Integrated pathogen load and dual transcriptome analysis of systemic host-pathogen interactions in severe malaria

One sentence summary: Combining simultaneous host and parasite RNA-sequencing and estimates of total body parasite load reveals distinct mechanisms associated with different aspects of severe malaria pathophysiology in humans.

Authors: Hyun Jae Lee[1], Athina Georgiadou[2], Michael Walther[3], Davis Nwakanma[3], Lindsay B. Stewart[4], Michael Levin[2], Thomas D. Otto[5#†], David J. Conway[4#], Lachlan J. Coin[1#], Aubrey J. Cunnington[2*]

Affiliations :

[1]Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia

[2]Section of Paediatrics, Imperial College, London, UK

[3] Medical Research Council Gambia Unit, Fajara, The Gambia

[4] Department of Pathogen Molecular Biology, London School of Hygiene and Tropical Medicine, UK

[5] Wellcome Trust Sanger Centre, Hinxton, UK

* To whom correspondence should be addressed: a.cunnington@imperial.ac.uk (AJC)

[#]Denotes equal contribution

†Current address: Centre of Immunobiology, Institute of Infection, Immunity & Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom

ThomasDan.Otto@glasgow.ac.uk

Abstract

The pathogenesis of infectious diseases depends on the interaction of host and pathogen but their simultaneous assessment is rare. In *Plasmodium falciparum* malaria, host and parasite processes can be assessed by dual RNA-sequencing of blood. Here we integrate dual transcriptome analyses with estimates of parasite load and detailed clinical information from 46 Gambian children to reveal mechanisms driving the systemic pathophysiology of severe malaria. We report hundreds of human and parasite genes differentially expressed between severe and uncomplicated malaria, with distinct profiles associated with coma, hyperlactatemia and thrombocytopenia. High expression of neutrophil granule-related genes was consistently associated with all severe malaria phenotypes. We observed previously uncharacterized, severity-associated variation in expression of parasite genes which determine cytoadhesion and rigidity of infected erythrocytes and parasite growth rate. Up to 99% of human differential expression was driven by differences in parasite load, whilst parasite gene expression showed little association with parasite load. Co-expression analyses revealed interactions between species, with prominent co-regulation of translation genes in severe malaria hinting at a molecular arms race between host and parasite. In multivariate analyses increased expression of granulopoiesis and interferon-γ related genes, together with inadequate suppression of type-1 interferon signalling best explained severity. Our findings highlight the importance of considering both host and pathogen when interpreting the pathogenesis of infectious diseases. We implicate simultaneous contributions of host and parasite genes to human severe malaria pathogenesis, identifying potential new targets for adjunctive therapy.

Introduction

Most studies of infectious disease pathogenesis focus on either host or pathogen, despite the fact that outcome is determined by their combined effects and interaction. Dual RNA-sequencing has developed as a method for transcriptomic assessment of such interactions (*1, 2*), although it has not been widely applied to study systemic infection in humans. Malaria is a good paradigm for this, because the pathogenic stage of the parasite is restricted to blood and within the blood many of the important interactions between parasites and host leukocytes occur (*3*). The blood is also the conduit for systemic responses to infection, and gene expression in blood will reflect the inflammatory and metabolic milieu to which leukocytes and pathogens are exposed.

Plasmodium falciparum malaria is one of the most important infectious diseases affecting humankind (*3*). Most malaria deaths occur in children, in whom three major severe malaria (SM) syndromes are associated with increased risk of death: cerebral malaria (CM, manifesting as coma), hyperlactatemia / acidosis (HL, often manifesting as deep breathing), and severe anemia (*3-6*). These are usually accompanied by the typical laboratory finding of low platelet count (thrombocytopenia) (*7*). SM syndromes can occur in isolation or in overlapping combinations (*4*), and mortality is highest when CM and HL coexist (CH) (*4*). SM is most likely when there is a high parasite load (*8-11*) and numerous accompanying pathophysiological derangements have been described (*4, 5, 12*), broadly arising from inflammation, vascular endothelial dysfunction and parasite sequestration (accumulation of parasitized erythrocytes in small blood vessels, obstructing blood flow) (*4, 13*). These mechanisms interact with one another so defining their individual contributions to specific features of SM is challenging (*13, 14*).

The host immune response is a major determinant of outcome in rodent models of SM (*15, 16*) and it has long been supposed that an excessive host response may also contribute to some forms of human SM (*17, 18*). A similar concept exists to explain severity in other infections such as bacterial sepsis, Ebola and Respiratory Syncytial Virus (*19-21*), however direct evidence is often lacking and the confounding effect of pathogen load on the magnitude of the host response is rarely quantified.

Controlled human infection models have provided insights into the relationship between pathogen load and early immune responses to infection (*22*) but cannot be extended to investigate associations with severity. To better understand the pathogenesis of severe infection an integrated view of host-pathogen interaction at a systemic scale and accounting for variation in pathogen load is required.

The feasibility of host and parasite dual-RNA sequencing in malaria has been demonstrated in individuals with uncomplicated malaria (UM) (*23*). Here, we extend the application of dual-RNA sequencing to infectious disease pathogenesis by integrating analysis of gene expression with detailed clinical and laboratory data which characterise the systemic pathophysiology of SM. We refine the analysis further by accounting for three major confounders which may vary within and between severity groups: leukocyte proportions in blood, parasite developmental stage distribution, and variation in whole body pathogen load. We characterise human and parasite differential gene expression between SM and UM, the association of gene expression with specific pathophysiological features, the role of parasite load in determining the host response, and networks of interaction between host and parasite gene expression. These data provide a unique insight into host-pathogen interactions associated with severity of infection in humans and reveal novel perspectives on the likely pathogenic mechanisms of human SM.

Results

Host and parasite RNA-sequencing

We performed dual-RNA sequencing on whole blood of 46 Gambian children with *P. falciparum* UM (n=21) and SM (n=25) (Table 1). These children were recruited from a region with relatively low malaria transmission and consistent with this (*4*) the SM group contained children with CM, HL and CH, but no cases of severe anemia. After exclusion of parasite *var*, *rifin*, *stevor* (*14*) and other highly polymorphic regions for which reference genome-based mapping is not possible (see Methods), we obtained medians of 26.6 million (26.6 million SM, 26.7 million UM, P=0.913) human and 9.61

million (10.3 million SM, 5.03 million UM, P=0.346) parasite uniquely mapped reads from each subject (Fig 1A). We detected expression of 12253 human and 3880 parasite genes. Commensurate with the high parasitemias seen in these children (Table 1), parasite read depth was considerably greater than in a previous study of Indonesians with UM (*23*).

Deconvolution and adjustment for cellular heterogeneity

Systemic infection provokes changes in blood leukocyte subpopulations which can dominate changes in gene expression and confound their interpretation (*24*). Amongst our study subjects there were significant differences between clinical groups in the proportions of neutrophils (*P*=0.01) and lymphocytes (*P*=0.05) in blood, although there were no significant differences in absolute cell counts (Table 1). To overcome this we performed gene signature-based deconvolution (*25*) to estimate heterogeneity in the contribution of the major leukocyte subpopulations to the RNA in our samples (Fig 1B, Supplementary Fig 1). Parasite gene expression *in vivo* is also dominated by the mixture of parasite developmental stages at the time of sampling because there is phasic variation in gene expression (*26*) and increasing total RNA content during the intraerythrocytic developmental cycle (*27*). Therefore we sought to apply the same approach with reference gene signatures derived from highly synchronous parasite cultures (*26, 28*) to identify the contribution of parasites at different developmental stages (Fig 1C). Since the method was developed and validated for distinct human cell types we confirmed the effectiveness of estimation of parasite developmental stage mixture by comparison with previously proposed stage-specific marker genes (*29*), and by assessment of performance in synthetic datasets of known composition (Supplementary Fig 2). We compared relative contributions of parasite developmental stages between SM and UM samples and observed a trend towards greater contributions from late stage asexual parasites and gametocytes in children with SM (Fig 1D), consistent with previous reports (*30*).

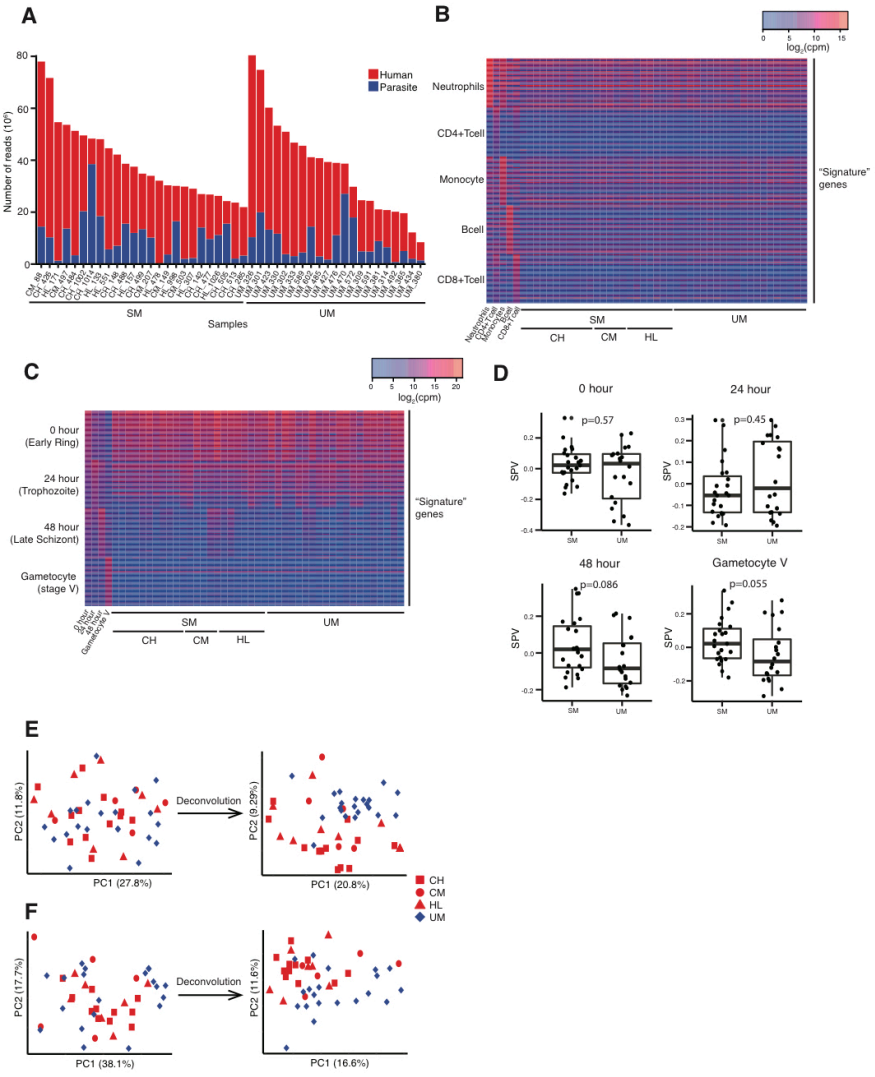


Figure 1. Whole blood dual RNA-sequencing and deconvolution. (**A**) Uniquely mapped reads from human (red) and *P. falciparum* (blue) from subjects with severe (SM, n=25) and uncomplicated malaria (UM, n=21). (**B,C**) Heatmaps showing signature gene expression for different leukocyte (**B**) and parasite developmental stage (**C**) populations (rows) and their relative intensity in individual subjects with SM, including different SM phenotypes (CH, cerebral malaria plus hyperlactatemia;

CM, cerebral malaria; HL, hyperlactatemia), and UM (columns). (**D**) Surrogate proportion variables for parasite developmental stages compared between SM and UM using the Mann-Whitney test (bold line, box and whiskers indicate median, interquartile range and 1.5-times interquartile range respectively). (**E,F**) Principal component plots showing the effect of deconvolution on the segregation of subjects with UM and SM, adjusting human (**E**) and parasite (**F**) gene expression for differences in proportions of leukocytes or parasite developmental stages respectively. Analyses of human gene expression (**B**,**E**): SM, n=25; UM, n=21. Analyses of parasite gene expression (**C**,**D**, **F**): SM, n=23; UM, n=20.

In order to remove the confounding effect of heterogeneity in leukocyte and parasite mixtures we adjusted gene expression values for cell mixtures, essentially allowing us to compare gene expression as if all subjects had the same leukocyte and parasite population compositions. Adjustment for heterogeneity in the mixture of leukocytes and parasite developmental stages improved segregation of SM and UM cases (Fig. 1E, F; multivariate ANOVA for human gene expression P=0.0013 and P = 0.00012, and for parasite gene expression P=0.0049 and P=0.00019, before and after adjustment respectively). Therefore we used adjusted gene expression values for all subsequent analyses.

Human differential gene expression in severe malaria

Considering all subjects together there were 770 human significantly differentially expressed genes (DEGs, with false discovery rate (FDR)-adjusted P<0.05) between SM and UM (Fig 2A, Supplementary Table 1). Amongst these some were highly significant and conspicuously higher expressed in SM relative to UM (high log-fold change), most notably *MMP8* (matrix metallopeptidase 8), *OLFM4* (olfactomedin 4), *DEFA3* (defensin A3), and *ELANE* (neutrophil elastase), all encoding neutrophil granule proteins (*31*). Due to the adjustment for cellular heterogeneity this likely reflects a true increase in transcription of these genes rather than a greater proportion of neutrophils in the blood. We performed Gene Ontology (GO) analyses to better understand the biological functions of the DEGs (Fig 2B) and identified enrichment of genes controlling co-translational protein targeting, cell motility and immune response functions (Supplementary Table 2). We used Ingenuity Pathway Analysis to predict upstream regulators of the differentially expressed genes, and colony stimulating factor 3 (CSF3, also known as granulocyte colony stimulating factor, GCSF), Fas cell surface death receptor, and Prostaglandin E receptor 2 signalling were the most significant (Supplementary Table 3). We repeated similar analyses to identify and interpret DEGs between UM and different clinical phenotypes of SM (Supplementary Fig 3, Supplementary Tables 1&2), which revealed many consistent associations but also some notable differences. For example the number of DEGs was

substantially increased when comparing the subgroup of subjects with CH (the most severe phenotype, n=12) vs UM, possibly reflecting their greater severity and homogeneity of disease (Supplementary Table 2, Supplementary Figure 3). The most highly expressed DEGs in CH relative to UM included neutrophil granule genes but also heat shock protein genes.

In order to gain a greater insight into specific pathophysiological processes we examined the quantitative association of gene expression with clinical and laboratory parameters which characterise specific aspects of SM pathophysiology, namely conscious level (using the Blantyre Coma Scale, BCS), blood lactate concentration, platelet count and hemoglobin concentration (Supplementary Table 4). 738 genes were associated with BCS (Supplementary Table 4) and decreasing conscious level (lower BCS) was associated with higher expression of genes involved in the cell cycle, and lower expression of genes involved in MHC class I antigen presentation and interferon gamma signalling (Fig 2B, Supplementary Table 2). Predicted upstream regulators included estrogen receptor 1 and transglutaminase 2 (Supplementary Table 3). 1012 human genes were significantly correlated with lactate concentration (Supplementary Table 4), amongst which immune response pathways were again prominent, but a negative association between lactate and type 1 IFN signalling was particularly notable (Fig 2B, Supplementary Table 2), and the most significant predicted upstream regulators were IFN-γ, IFN-α, and TNF (Supplementary Table 3). 178 genes were significantly associated with platelet count (Supplementary Table 4) and the most enriched pathways differed considerably from those in the preceding analyses (Fig 2B, Supplementary Table 2), with nucleosome assembly (predominantly histone genes), coagulation, and response to wounding genes, all negatively correlated, and the most significant predicted upstream regulators being IL13, RB transcriptional corepressor 1, and IL1RN (Supplementary Table 3). No human genes were significantly correlated with hemoglobin. Taken together these results identify common transcriptional features of SM but also implicate distinct mechanisms underlying the different pathophysiological processes which can occur in SM.

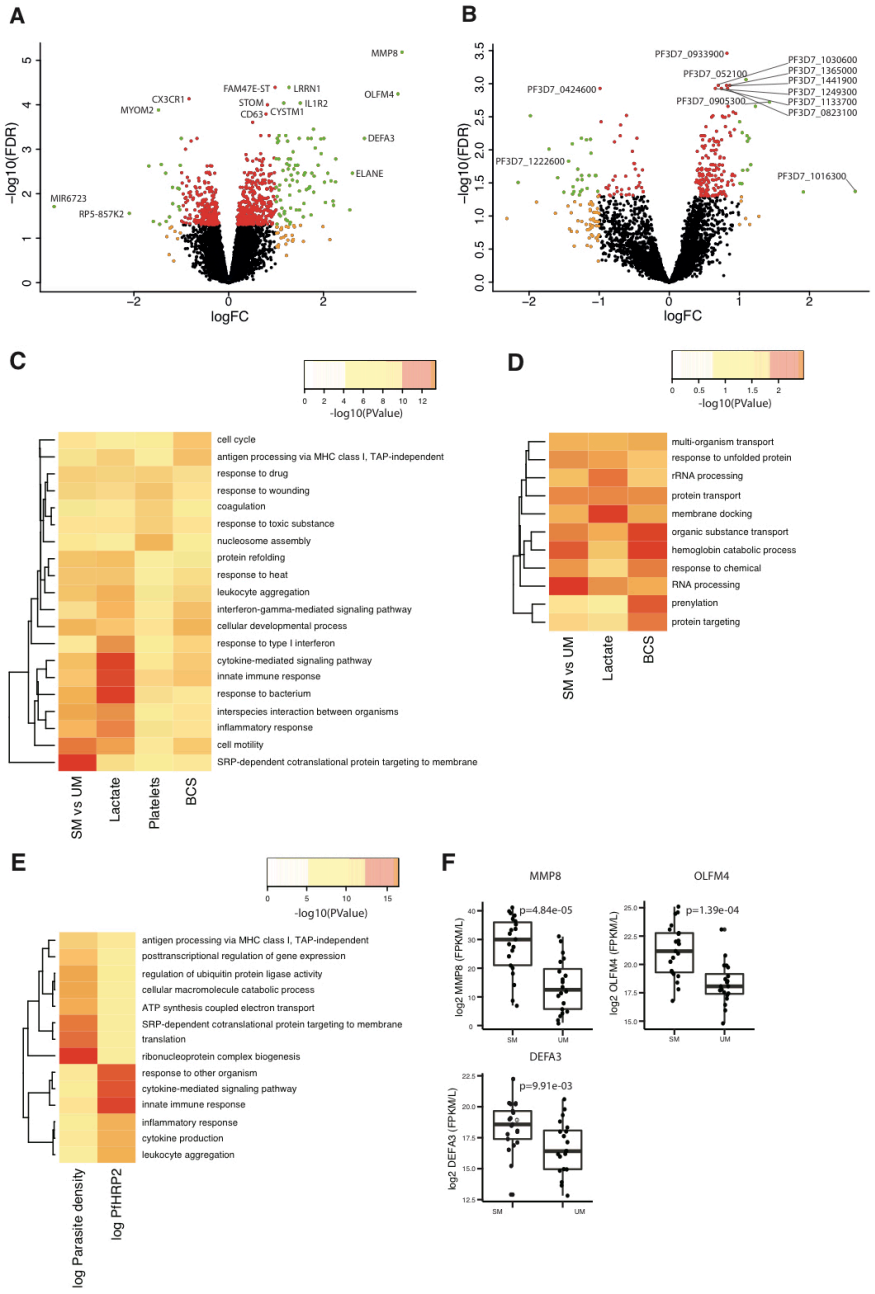

A
B
C
D
E
F
−log10(FDR)
logFC
MMP8
OLFM4
DEFA3
ELANE
CX3CR1
MYOM2
FAM47E-ST
LRRN1
STOM
IL1R2
CD63
CYSTM1
MIR6723
RP5-857K2
PF3D7_0933900
PF3D7_0424600
PF3D7_1222600
PF3D7_052100
PF3D7_0905300
PF3D7_1030600
PF3D7_1365000
PF3D7_1441900
PF3D7_1249300
PF3D7_1133700
PF3D7_0823100
PF3D7_1016300
-log10(PValue)
cell cycle
antigen processing via MHC class I, TAP-independent
response to drug
response to wounding
coagulation
response to toxic substance
nucleosome assembly
protein refolding
response to heat
leukocyte aggregation
interferon-gamma-mediated signaling pathway
cellular developmental process
response to type I interferon
cytokine-mediated signaling pathway
innate immune response
response to bacterium
interspecies interaction between organisms
inflammatory response
cell motility
SRP-dependent cotranslational protein targeting to membrane
SM vs UM
Lactate
Platelets
BCS
multi-organism transport
response to unfolded protein
rRNA processing
protein transport
membrane docking
organic substance transport
hemoglobin catabolic process
response to chemical
RNA processing
prenylation
protein targeting
antigen processing via MHC class I, TAP-independent
posttranscriptional regulation of gene expression
regulation of ubiquitin protein ligase activity
cellular macromolecule catabolic process
ATP synthesis coupled electron transport
SRP-dependent cotranslational protein targeting to membrane
translation
ribonucleoprotein complex biogenesis
response to other organism
cytokine-mediated signaling pathway
innate immune response
inflammatory response
cytokine production
leukocyte aggregation
log Parasite density
log PHRP2
p=4.84e-05
p=1.39e-04
p=9.91e-03
log2 MMP8 (FPKM/L)
log2 OLFM4 (FPKM/L)
log2 DEFA3 (FPKM/L)
SM
UM

Figure 2. Association of gene expression with features of severe malaria and parasite load. (**A**) Volcano plot showing extent and significance of up- or down- regulation of human gene expression in SM compared with UM (red and green, P <0.05 after Benjamini-Hochberg adjustment for false discovery rate (FDR); orange and green, absolute log_2-fold change (FC) in expression > 1; SM n=25, UM=21). (**B**) Heatmap comparing enrichment of gene ontology terms for human genes significantly differentially expressed between SM and UM or significantly associated with blood lactate, platelet count or Blantyre Coma Scale (BCS). (**C**) *P. falciparum* differential gene expression in SM compared to UM (colour coding as in (**A**), SM n=23, UM=20). (**D**) Heatmap comparing enrichment of gene ontology terms for parasite genes significantly differentially expressed between SM and UM and or significantly associated with blood lactate or BCS. (**E**) Heatmap comparing gene ontology terms for human genes significantly associated with log parasite density and log PfHRP2. (**F**) Comparison of selected neutrophil-related gene expression multiplied by absolute neutrophil count in blood between SM (n=) and UM (n=) (Wilcoxon)

Parasite differential gene expression in severe malaria

There were 236 parasite DEGs between SM and UM (Fig 2C, Supplementary Table 5). The parasite DEG with highest expression in SM relative to UM was *PF3D7_1016300* (Supplementary Table 3) which encodes a glycophorin binding protein (GBP) expressed in the cytoplasm of infected erythrocytes and which influences adhesion and rigidity of the red cell (*32, 33*). The most down-regulated parasite DEG (with known function) in SM relative to UM was *PF3D7_1222600*, which encodes the AP2 domain transcription factor AP2-G. This protein controls the balance between gametocytogenesis and asexual replication, and knockout of the orthologue in *P. berghei* ANKA significantly enhances *in vivo* asexual parasite growth rate (*34, 35*). As with human gene expression there were similarities in patterns of DEGs in the comparison of different clinical phenotypes of SM vs UM, and again the most DEGs were found in the comparison with the CH group (Supplementary Fig 3, Supplementary Table 5). Here the most differentially expressed parasite genes included *PF3D7_0202000* (knob-associated histidine rich protein), *PF3D7_1016300* (GBP), *PF3D7_0201900* (erythrocyte membrane protein 3), and *PF3D7_0424600* (PHIST-b protein), all of which encode proteins which interact with the erythrocyte cytoskeleton to influence cytoadhesion and deformability of the infected erythrocyte, making them highly plausible determinants of severity (*33*). Parasite DEGs in SM compared to UM were enriched in specific biological functions including RNA processing, protein transport, and hemoglobin catabolism (Fig 2D, Supplementary Table 6). 445 parasite genes were significantly associated with BCS, and those most significantly correlated with decreasing conscious level were *PF3D7_0919800* (TLD domain-containing protein), *PF3D7_1133700* (FHA domain-containing protein), and *PF3D7_1408200* (AP2 domain transcription factor AP2-G2), the latter two both important determinants of asexual parasite growth rate (*35, 36*) (Supplementary Table 7). The most enriched functions associated with BCS included transport, hemoglobin catabolism and prenylation (Fig 2D). 100 parasite genes were significantly correlated with lactate (Supplementary Table 7). The most significant was *PF3D7_0201900* (encoding erythrocyte membrane protein 3, EMP3), consistent with infected erythrocyte rigidity and

cytoadhesion (*32, 37, 38*) being important determinants of microvascular obstruction and hyperlactatemia (*39*). Pathway enrichments amongst these genes differed from those most associated with BCS and included membrane docking and rRNA processing (Fig 2D and Supplementary Table 6). Few parasite genes were associated with hemoglobin concentration or platelet count (Supplementary Table 7). Taken together, these findings indicate that different patterns of parasite gene expression are associated with, and may therefore contribute to, specific aspects of host pathophysiology.

In order to establish whether changes in parasite gene expression might be cause or consequence of SM, we tested the effect of hyperlactatemia on parasite gene expression *in vitro*. 61 genes were differentially expressed between lactate supplemented (n=4) and control (n=5) early ring-stage parasite cultures, particularly enriched in genes associated with transcription and RNA processing (Supplementary Tables 5 and 6). Surprisingly two of the genes most highly induced by lactate supplementation were *PF3D7_1016300* (GBP) and *PF3D7_0202000* (Knob-associated histidine rich protein), genes which were also highly expressed in the CH phenotype. This suggests lactate may influence the virulence phenotype of parasites, consistent with another recent report that *Plasmodium* can sense and respond to the host metabolic environment (*40*).

Effect of parasite load on host and parasite gene expression

Previous studies have shown a correlation between the expression levels of host genes and circulating parasitemia (*41, 42*). There were significant differences in parasite load between our subjects with SM and UM (Table 1) and so we were interested to determine the extent to which this explained the differences in whole blood gene expression. Peripheral blood parasite quantification (parasite density) underestimates the total number of parasites in the body because of sequestration of parasites in the microvasculature (*13, 14*). The soluble parasite protein, *P. falciparum* histidine rich protein 2 (PfHRP2), has been used as a plasma biomarker of total parasite load (circulating plus sequestered) and is more strongly associated with severity (*8, 9, 11*) and death

(*8, 9*). Therefore we examined the association of host and parasite gene expression with both circulating parasite density and PfHRP2 (restricting comparisons to subjects with data for both). We found 1886 human genes significantly (FDR P<0.05) correlated with log parasite density and 616 significantly correlated with log PfHRP2 (102 common to both), whilst only 2 and 10 parasite genes were significant in the corresponding analyses (none common to both) (Supplementary Tables 4 and 7). Human genes correlated with log parasite density were particularly enriched in pathways related to translation (especially exported proteins), oxidative phosphorylation, and antigen presentation (Fig 2E, Supplementary Table 2), with predicted upstream regulation by RPTOR independent companion of MTOR complex 2 (RICTOR), hepatocyte nuclear factor 4 alpha (HNF4A) and X-box binding protein 1(XBP1); Supplementary Table 3). Genes correlated with log PfHRP2 were particularly enriched in innate immune response functions (Fig2E, Supplementary Table 2), with predicted regulation by interferon-γ (IFN-γ), transglutaminase 2, and IFN-α2 (Supplementary Table 3). These findings suggest that the nature of the systemic host response is associated with the localisation of parasites.

We next asked to what extent the differences in gene expression between SM and UM phenotypes were dependent on parasite load. Restricting analyses to subjects with both parasite density and PfHRP2 measurements, the number of human SM vs UM DEGs remained almost unchanged after adjustment for parasite density but was reduced by 98.6% after adjustment for PfHRP2, whilst parasite DEGs changed much less after either of the same adjustments (Table 2, Supplementary Tables 1 and 5). Findings were similar when adjusting for parasite load in comparisons of each of the SM subtypes vs UM (Supplementary Tables 1 and 5). Repeating this analysis in an independent dataset of human microarray gene expression in Malawian children with CM (*43*), revealed 994 DEGs (FDR P<0.05) between children with (n=55) and without (n=17) malaria-associated retinopathy (Supplementary Table 1), 608 (61%) DEGs after adjustment for parasite density, and no DEGs after adjustment for PfHRP2.

277 Taken together the preceding findings suggest that total body parasite load, as represented by
278 PfHRP2, is a dominant determinant of host gene expression in malaria, particularly inflammatory and
279 immune response genes, and differences in total body parasite load drive the majority of the human
280 gene expression differences between SM and UM. However, if genes remain associated with severity
281 after adjustment for parasite load this may indicate intrinsic variation in the host response which
282 determines susceptibility to severe disease. In our dataset, only 13 genes remained significant after
283 adjustment for PfHRP2 (Table 2, Supplementary Table 1). Of particular interest amongst these,
284 MMP8 (also known as collagenase 1) is a metallopeptidase which causes endothelial barrier damage
285 in several infection models (*44, 45*); *AZI2* (also known as NF-Kappa-B-Activating Kinase-Associated
286 Protein 1, NAP1) encodes a regulator of the type 1 interferon response (*46*), a pathway which is
287 known to control severity of disease in rodent malaria models (*47*); whilst CX3CR1 is the receptor for
288 fractalkine (a biomarker of CM in humans (*48*)),is expressed on subset of monocytes which are
289 particularly efficient at killing malaria parasites (*49*), and controls the trafficking of monocytes during
290 inflammation (*50*).
291 Parasite load was also a major driver of the associations between human gene expression and BCS,
292 lactate, and platelet count, but its importance did vary between them (Table 2). Interestingly the few
293 remaining genes significantly associated with lactate after adjustment for PfHRP2 included *PKM*
294 (encoding the glycolytic enzyme pyruvate kinase M) and *GYS1* (encoding the glycogenic enzyme
295 glycogen synthase 1) (Supplementary Table 4), suggesting hyperlactatemia is partly associated with
296 parasite load-independent variation in control of host glucose metabolism. The number of genes and
297 biological functions associated with platelet count were notably less dependent on parasite load.
298
299 *Co-expression networks of host and parasite genes*
300 Whilst independent analyses of host and parasite gene expression associations with severity are
301 enlightening, dual-RNA sequencing can also be used to identify molecular interactions within and
302 between species and their associations with severity (*1*). Expression of groups of genes with

common functional roles are often highly correlated and can be identified through co-expression network analysis (*51*). We applied this methodology to the paired host and parasite gene expression from each individual, to identify groups of genes originating from either or both species, on the basis of their correlated expression. We term these groups of genes "modules", and we named each module according to the "hub gene" which has the greatest connectivity to other genes within the module. First we analysed all subjects together and a network with 26 modules was generated (Fig 3 and Supplementary Table 8): 10 containing exclusively human genes, 5 exclusively parasite genes, and 11 with both human and parasite genes (most of these highly skewed to a single species). Only the HSPH1 (heat shock protein family H (Hsp110) member 1) module contained more than 10 genes from both human and parasite, strongly enriched in human heat shock response genes and parasite RNA metabolism genes. All modules showed significant functional enrichments regardless of host or parasite origin. The composite expression of genes within a module can be described by a module eigengene value (*52*) and there were significant associations between module eigengene values and severity, parasite load, BCS and laboratory parameters (Figure 3). Some host-dominated and parasite-dominated modules were also highly correlated with each other, most notably the *RPL24* (ribosomal protein L24) module (highly enriched in translation pathways) was strongly correlated with the remarkably homologous *PF3D7_0721600* (putative 40S ribosomal protein S5) parasite module. We excluded read mapping errors as an explanation for this, and suggest that this indicates co-regulation of conserved host and parasite translation machinery. Furthermore, most of these genes were also differentially expressed between SM and UM. This association may indicate a molecular arms race between parasite and host to synthesise proteins for defence and survival, respectively. In excess, these proteins may contribute to tissue damage in SM.

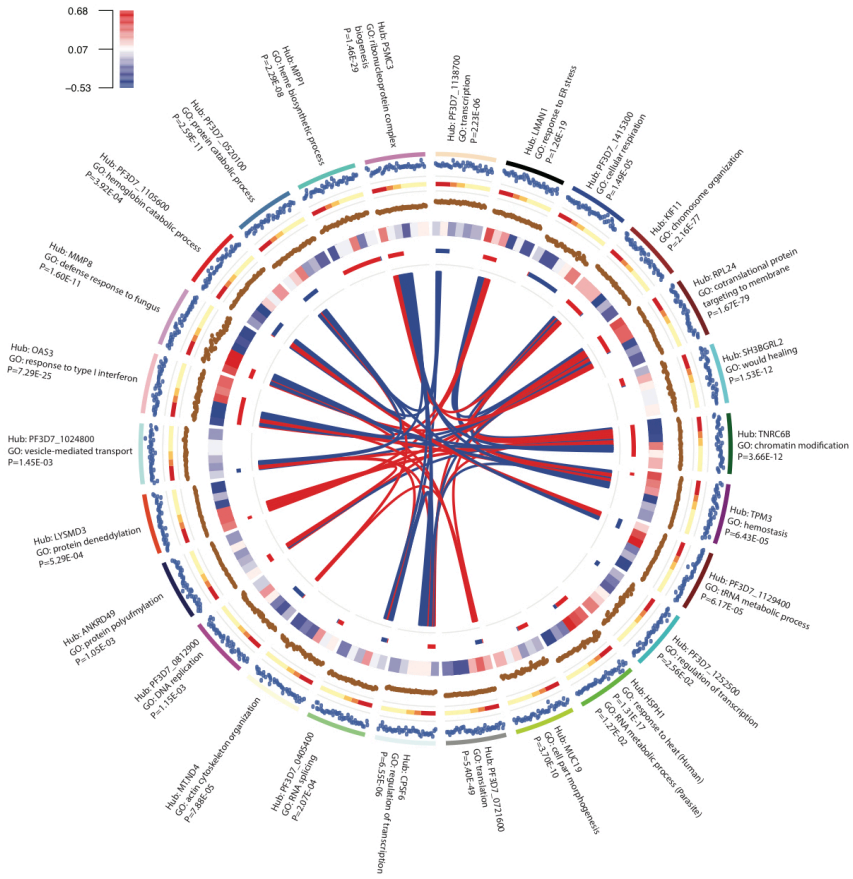


Figure 3. Interspecies gene expression modules and their associations with severity. Circos plot showing gene expression modules obtained from whole genome correlation network analysis using expression of all human and parasite genes from each subject (SM, n=22; UM, n=19) as the input. From outside to inside: labels, hub gene and most enriched GO term (with enrichment P-value) for each module; track 1, module eigengene value for each subject; track 2, clinical phenotype (Red=CH, Orange=CM, Green=HL, Yellow=UM); track 3, hub gene expression (log CPM) for each subject; track 4, heatmap for correlation with laboratory measurements (clockwise, blocks: log parasite density, log PfHRP2, lactate, platelets, hemoglobin, BCS; colour intensity represents correlation coefficient as shown in colour key); track 5, module size and composition (length proportional to number of genes

in module; red, human genes; blue, parasite genes); polygons connect modules with significant (FDR P<0.01) Pearson correlation between eigengene values (width proportional to $-\log_{10}$ FDR P-value; red=positive correlation, blue=negative correlation)

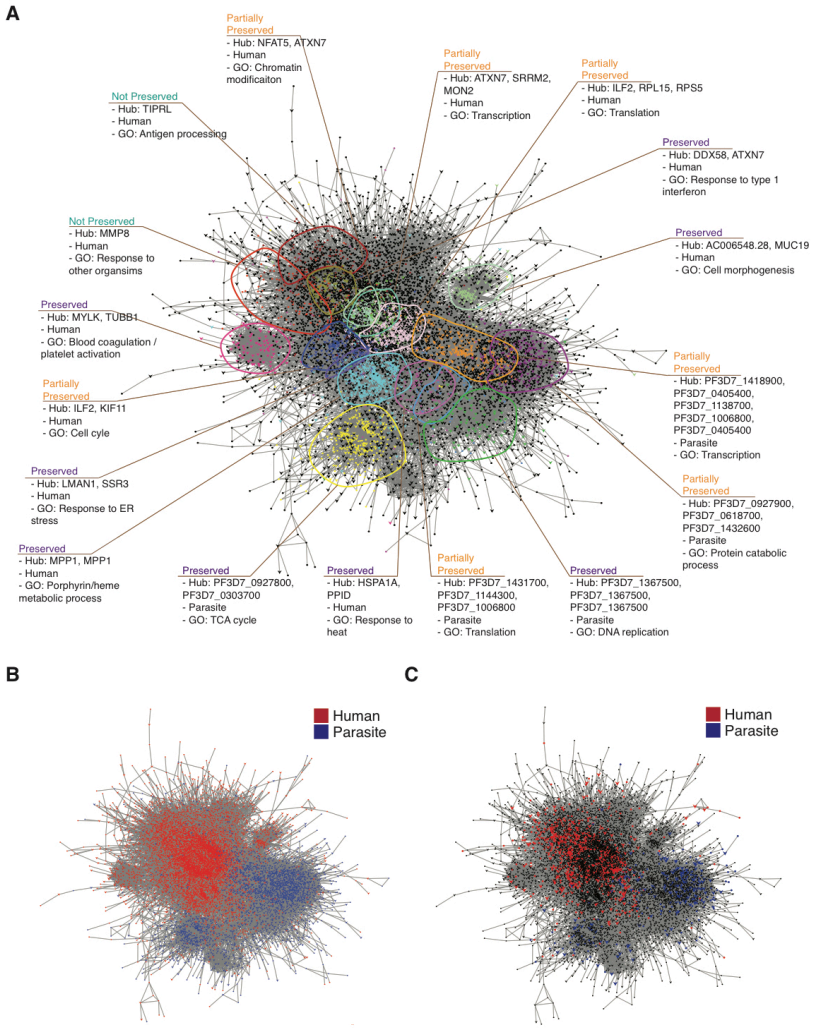
Association of co-expression modules with severity

Co-expression network modules can be used as units of analysis, affording considerable dimension reduction for whole-genome expression data. We used module eigengene values (*51, 52*) and parasite load (with which many modules were correlated, Fig 3) in linear regression models to determine the best within-sample predictors of severity, starting with all significant univariate associations and proceeding by backward selection (Supplementary Table 9). The best multivariate model combined *MMP8*, *OAS1* (2'-5'-oligoadenylate synthetase 1) and *LYSMD3* (LysM, putative peptidoglycan-binding, domain containing 3) module eigengenes, but not parasite load. Interestingly these modules represent distinct aspects of the immune response: the *MMP8* module, highly enriched in defence response genes with predicted upstream regulators CEBPA (CCAAT/enhancer binding protein alpha, a myeloid transcription factor) and CSF3, likely reflects granulopoiesis (*31*); the *OAS1* module is highly enriched for type 1 IFN response genes; the small *LYSMD3* module, with limited GO enrichment, contains a functional network around IFN-γ (Supplementary Figure 4). The direction of association of the *OAS1* module with severity changed from negative in univariate analysis to positive in the multivariate analysis, suggesting that inadequate suppression of the type-1 IFN response in conjunction with upregulation of granulopoiesis and IFN-γ signalling may contribute to pathogenesis.

Differential co-expression in severe malaria

Considering all subjects together for generation of co-expression networks maximises power to detect consistently co-regulated genes but may not identify sets of genes where co-regulation is altered by severity. For this reason we also created separate co-expression networks for UM and SM and compared the modules to identify differential co-expression (Fig 4, Supplementary Table 10). Eight modules showed significant preservation between networks, seven were partially preserved, and two were unique to SM (Figure 4A). Partial preservation was common amongst modules comprised predominantly from human or parasite genes (Figure 4A,B), and module preservation was

not dependent on the proportion of module genes differentially expressed between SM and UM (Figure 4A,C). A *MMP8* module was identified (exclusively human genes, many encoding neutrophil granule and phagosome components) uniquely in SM subjects, and 38% of its member genes were DEGs in the comparison between SM and UM. The module was enriched in host defence functions (Supplementary Table 10) and predicted to be regulated by CEBPA, CSF3 and TNF. These findings strongly suggest the MMP8 module represents emergency granulopoiesis (*31*) and mark this as a specific feature of SM. The *TIPRL* (TOR Signaling Pathway Regulator) module (99.2% human genes) was also unique to SM but contained very few (1.3%) DEGs, had limited GO enrichment (Supplementary Table 10), and the most significant predicted upstream regulator was the transcription factor HNF4A. Both TIPRL and HNF4A have regulatory roles in metabolic, inflammatory and apoptosis signal pathways, so the minimal change in expression of this module may represent an aberrant response in SM (*53, 54*). Amongst the partially preserved modules we observed that host and parasite translation pathways were more tightly co-regulated in SM than UM, genes being distributed across fewer modules in SM (Fig 4A, Supplementary Table 10). This once again suggests that there is an interaction between these processes that is associated with severity.



383

384 **Figure 4. Severity-associated differential co-expression within the interspecies gene expression**

385 **network.** (**A-C**) Cytoscape visualisation of merged co-expression networks derived separately from

386 SM (n=22) and UM (n=19). Networks were merged such that genes found in both sub-networks

387 (represented as arrow-shaped, larger-sized nodes) are connected to genes found in only one sub-

388 network (represented as circular-shaped and smaller-sized nodes). (**A**) Genes and gene clusters are

389 coloured and annotated by module, species, most enriched gene ontology terms, and conservation

between sub-networks (preserved, module pairs from SM and UM sub-networks display highly significant overlap with each other and much less significant overlap with other modules; partially preserved, module clusters display significant overlaps with two or more modules in the other sub-network; unique, gene clustering only found in one sub-network); genes in black do not belong to any characterized module. (**B**) Identical network layout with genes coloured by species (red, human; blue, *P. falciparum*). (**C**) Identical network layout with genes coloured by whether they are significantly differentially expressed in SM vs UM (red, human; blue, *P. falciparum*; black, not differentially expressed).

Interaction of parasite and host gene expression accounting for parasite load

We sought to integrate pathogen load into analysis of interaction between host and parasite gene expression. To reduce dimensionality we generated human-only gene expression modules from all subjects (Supplementary Table 11), identified those modules significantly associated with severity, and then identified parasite genes with significant (FDR *P* <0.05) pairwise associations with these modules in a linear model accounting for log PfHRP2. The human modules associated with severity were similar to those identified in preceding analyses (Table 3). The most significantly associated parasite genes and the most enriched parasite GO terms were those involved in RNA processing and translation (Table 3, Supplementary Table 11), suggesting that these processes in the parasite drive multiple aspects of the host transcriptional response independent of their effect on parasite load.

Neutrophil-related proteins in plasma

The most differentially expressed genes in comparisons between SM and UM encode neutrophil granule proteins. Relationships between transcription, translation, storage and release of granule proteins are expected to be complex, but we sought evidence of correlation between gene expression and circulating levels. In subjects with residual stored plasma we found significant correlations between gene expression and plasma concentrations of defensin A3 (P = 0.0049, rho =0.47, n=34) and elastase (P = 0.045, rho 0.35, n=34) (Supplementary Figure 5). MMP8 expression was significantly correlated with plasma levels in subjects with SM (P = 0.02, rho = 0.59 , n=15), but not in UM (P = 0.37, rho = -0.21, n=20) or all subjects combined (P=0.88 , rho=-0.026, n=35). GCSF was predicted as a major upstream regulator of genes in the MMP8 module. We found that plasma GCSF concentrations were significantly correlated with the eigengene values for this module (P = 0.0030, rho 0.64, n=19). We tested whether neutrophil degranulation occurred in response to parasite material by stimulating healthy donor blood cells with *P. falciparum* schizont lysate, and detected significant increases in MMP8 release, reaching similar concentrations to those observed in plasma during malaria (Supplementary Figure 5).

Discussion

We used dual-RNA sequencing to identify simultaneous host and parasite gene expression and their systemic interactions associated with severity of *P. falciparum* malaria in humans. Our findings establish this as a paradigm for integrated understanding of infectious diseases and make a strong case that neither host nor pathogen should be studied in isolation.

We have identified many plausible associations between gene expression and features of severity, providing an unbiased insight into the pathogenesis of SM (Figure 5). One of our most striking findings was the overriding effect of parasite load on differences in human gene expression between SM and UM. Previous studies have examined the association between human gene expression and circulating parasitemia (*23, 41, 42*), but we found that estimation of total body parasite load was necessary to appreciate the full effect on host response. Our findings imply that the host response in SM is not excessive *per se*, but rather that it is an appropriate host response to an excessive pathogen load. This has important implications for malaria research and likely other infectious disease, immunology, and pathogenesis research in humans. Without accounting for pathogen load, associations between host factors (such as genetic variants or comorbidities) and severity of infection may well be misinterpreted. Unfortunately, total body pathogen load is much harder to measure in other infections in humans where pathogens are not restricted to the blood (*55*).

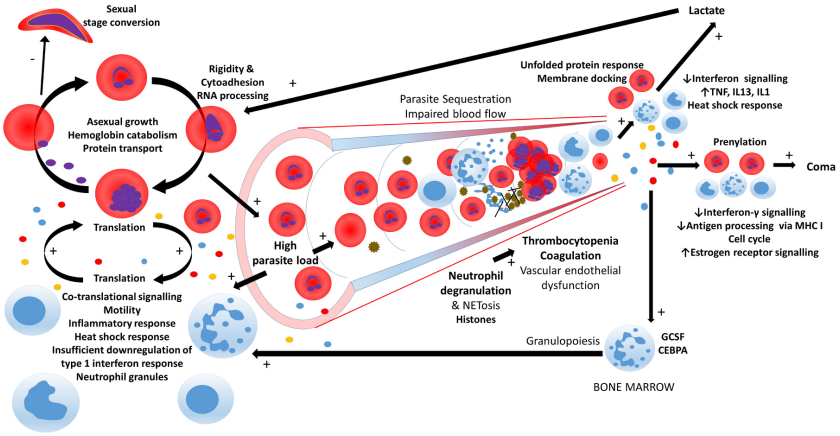


Figure 5. Host pathogen-interactions in severe malaria revealed through dual RNA-sequencing. Selected biological processes are illustrated, representing human and parasite gene expression associated with SM. High parasite load is the major driver of the human whole blood transcriptional response. Parasite genes regulating asexual growth rate, sexual differentiation, and biophysical properties of the erythrocyte associate with SM. Translation pathway gene expression in parasites correlates with expression of similar genes in human cells, suggesting that a "molecular arms race". Sequestration of high concentrations of cytoadherent and rigid parasitized eryrthrocytes in small blood vessels likely stimulates processes such as neutrophil degranulation, NETosis, and vascular endothelial activation, which then drive coagulation and thrombocytopenia, worsening microvascular dysfunction, and proinflammatory responses. GCSF promotes granulopoiesis and neutrophil granule production, potentially amplifying the pathological process. Subsets of host and parasite genes associated with either hyperlactatemia or coma point to differing pathophysiology of these clinical manifestations. For example, prenylation may enhance erythrocyte membrane presentation of parasite proteins needed for sequestration in the brain vasculature, while proinflammatory signalling may drive hyperlactatemia through aerobic glycolysis in immune cells. Parasites may respond to changes in lactate by altering gene expression to evade the host response.

460 Bold text identifies biological processes supported by the present study, non-bold text identifies

461 processes which are known to occur in malaria but not specifically investigated in the present study.

462

We found that specific sets of host and parasite genes were associated with different pathophysiological consequences of malaria, albeit that power to detect associations was limited in the smallest subgroup analyses. Distinct sets of host genes were correlated with BCS, lactate concentration and platelet count. Hyperlactatemia in malaria is often ascribed to anaerobic metabolism arising from microvascular obstruction by adherent and rigid parasitized erythrocytes (*4, 39*). These properties are partly determined by the expression of particular members of the *var*, *rif* and *stevor* gene families (*14*) , which we did not include in our analysis because the extreme degree of polymorphism prevents reference genome-based quantification. Despite this limitation we still found associations between lactate concentration and severity and *in vivo* variation in the expression of other parasite genes known to modify the biophysical properties of the infected erythrocyte. Some variation in parasite gene expression may have a genetic basis (*56*) but our *in vitro* data suggests that it may also occur as a response to the host environment.

The human genes most correlated with lactate were immune response-related, suggesting that inflammation and perhaps its effect on glycolysis may also be involved (*57, 58*). If lactate production associated with the strength of the host response then the changes in expression of parasite genes which might favour sequestration and evasion of innate immune cells in response to lactate exposure could be viewed as a parasite survival strategy. Cerebral malaria in humans is usually ascribed to parasite sequestration in the cerebral microvasculature, but the association between BCS and antigen presentation via MHC class 1, interferon-γ, and type-1 interferon signalling, would be consistent with the localisation of these immunopathological mechanisms away from the blood and into the brain microvasculature as seen in rodent experimental cerebral malaria (*15, 16, 59*).

Thromobocytopenia is almost invariable in malaria and its mechanism is poorly understood (*7*). Our findings implicate the well-recognised activation of endothelial surfaces and coagulation pathways in malaria (*7*) as a cause, but also lead us to suspect a role for histone-induced thrombocytopenia (*60, 61*).

Many of our results converge on a putative role for neutrophils in SM. Different analytical approaches repeatedly identified the association of genes encoding neutrophil granules (such as *MMP8, OLFM4*, *DEFA3*, *ELANE*) and upstream regulators of granulopoiesis (CSF3 and CEBPA) with severe outcomes. Granule proteins are enriched in immature neutrophils (*31, 62*) which are mobilised from the bone marrow to the circulation in malaria (*63, 64*), and there is plentiful evidence that neutrophil degranulation occurs in SM (*43, 65*). Release of neutrophil granule proteins can be highly damaging to host tissues (*62*), and increased production and release of these proteins could contribute to many of the pathological features of SM. Both neutrophil granule proteins and histones are released into the circulation during production of neutrophil extracellular traps (NETs) (*62, 66*), a phenomenon which has been described in malaria (*64*). It is noteworthy that similar neutrophil-related signatures are not reported in the whole blood transcriptomes of rodent models which have been examined to date (*42, 67*), creating a challenge for experimental validation. However neutrophil depletion has been shown to prevent experimental cerebral malaria (*68*), and whilst this is not a viable therapeutic option in humans, pharmacological inhibitors of specific neutrophil functions such as NETosis are being evaluated (*66*).

We observed an intriguing relationship between type-1 IFN responses and severity, which may help to tie together data from previous observations in humans and animal models. In a small study, expression of type-1 IFN response genes in blood from UM was higher than in SM, leading to the suggestion that this may be protective against developing SM (*69*). However we found that type-1 IFN response genes were negatively correlated with parasite load, suggesting down-regulation with increasing parasite load (and severity) is a more likely explanation. When we performed multivariate analyses using gene expression modules to explain severity, our results suggested that insufficient suppression of type-1 IFN signalling was in fact associated with severity. This would be more consistent with results in several animal models where genetic or antibody-mediated ablation of type-1 IFN signalling improves outcome (*70-73*).

513 Very few parasite genes correlated significantly with parasite load at the time of clinical presentation
514 which may be a consequence of its dynamic nature, determined by parasite growth rate, the
515 number of replication cycles in the host (duration of infection), and its reciprocal interaction the
516 constraining host response that it stimulates. Thus lower expression of the gene encoding ApiAP2-G
517 in SM may well cause increased asexual parasite growth rate (*34*) and make SM more likely, without
518 this gene exhibiting any significant correlation with parasite load. It may also seem paradoxical that
519 this gene is down-regulated in SM when gametocytes are usually more common in SM (*30*), a trend
520 we observed in our data. Development of mature gametocytes takes 10-12 days, for much of which
521 they are not in the systemic circulation (*74*), so we speculate that the mature gametocytes detected
522 at the time of clinical presentation may reflect preferential gametocytogenesis early in infection and
523 perhaps a subsequent reduction in ApiAP2-G promotes enhanced asexual replication and severe
524 disease.
525 We noted that both human and parasite translation pathways were associated with SM, and these
526 pathways showed the strongest evidence of interaction between species, with co-regulation
527 appearing tighter in more severe disease. Increased translation is important for production of host
528 defence effector proteins (*75*) and parasite proteins which enable survival (*76*) and it is feasible that
529 these processes drive each other. This raises the intriguing question of whether addition of a
530 translation inhibiting anti-malarial such as mefloquine (*77*) to standard artesunate treatment may
531 have added benefit in SM.
532 The differences in human and parasite gene expression between SM and UM were much clearer
533 after adjusting for heterogeneity of leukocyte population and parasite developmental stage.
534 Although the importance of accounting for such variation is well recognised (*24*), it is rarely done in
535 infectious disease transcriptomic studies. Several studies have used alternative methods to account
536 for parasite developmental stage distribution and have shown that this has a major impact on
537 observed associations between parasite gene expression and clinical phenotype (*78, 79*). The data
538 we have generated and comprehensive analyses we have performed following these adjustments

provide a unique and valuable resource for the research community. Whilst we cannot establish causation from an observational study such as this, our findings should be launch points for future work assessing the implicated mechanisms and their potential as targets for adjunctive therapies. The identification of multiple and sometimes distinct host and parasite mechanisms associated with differing aspects of pathophysiology potentially bodes ill for adjunctive therapies, which might need to have multiple targets and perhaps be personalized to differing SM manifestations. However the common association of neutrophil granule protein genes with all SM manifestations suggests that targeting neutrophil function may be a therapeutic strategy in the future.

Materials and Methods

Experimental design

The primary aim of the study was to analyse differential human and parasite gene expression between children with SM and UM and to determine its association with parasite load. Secondary aims were 1) to analyse differences in gene expression associated with different severe malaria syndromes and with continuous variables which are markers of pathophysiology, 2) to evaluate co-expression of host and parasite genes, and 3) to evaluate differential co-expression associated with SM. Sample size was determined pragmatically based on the availability of suitable samples with necessary clinical and laboratory data, since there is no established method to determine optimal sample size for dual RNA-Seq studies. We aimed to achieve close to 30 million mapped human reads and 5 million mapped parasite reads per sample and calculated the likely number of reads required to achieve this based on the percentage parasitemia and the likely amount of RNA per ring-stage parasite (*27*). From available RNA samples we aimed to have roughly equal numbers of severe and uncomplicated malaria cases and within the severe malaria cases we aimed to have eight subjects with each phenotype. After assessment of RNA quantity and quality some samples were unsuitable for RNA-sequencing which resulted in the final composition of groups being slightly unbalanced.

Subjects and samples

Gambian children (under 16 years old) with *P. falciparum* malaria were recruited from three peri-urban health centres, The MRC Gate Clinic, Brikama Health Centre, and The Jammeh Foundation for Peace Hospital, Serekunda, as part of a larger study of severe malaria (*11, 80, 81*). Informed consent was obtained from the child's parent or legal guardian for collection and subsequent use of samples. The study was approved by the Gambian Government / MRC Laboratories Joint Ethics Committee. All children underwent full clinical examination and were managed in accordance with the Gambian government guidelines. Malaria was defined by the occurrence of fever in the last 48 hours before recruitment and >5000 asexual parasites/µL in the peripheral blood. Subjects were further categorized into different severe malaria phenotypes using modified World Health Organization

criteria: cerebral malaria (CM) was defined as Blantyre Coma Score (BCS) of 1 or 2, or a BCS of 3 if the motor response was 1, AND no hypoglycaemia, no rapid improvement in response to fluid resuscitation, no suspicion of meningitis; hyperlactatemia (HL), blood lactate concentration > 5mmol/L; both CM and HL (CH) (*11*). At the time of presentation to the clinic, prior to any antimalarial treatment or blood transfusion, capillary blood was used for measurement of lactate and glucose concentrations and thick and thin blood films, venous blood was collected into EDTA for sickle cell screen and full blood count, PAXgene blood RNA tube (BD), and sodium heparin (BD) for plasma separation (*80*). Parasitemia was calculated using 50 high power fields on Giemsa-stained thin blood smears. Plasma *P. falciparum* histidine-rich protein II (PfHRP2) was measured by ELISA (Cellabs) (*11*). Full blood counts, including hemoglobin concentration, platelet count, total white cell count and three part differential (neutrophil, lymphocyte and monocyte counts) were made using a Medonic™ hematology analyser (Clinical Diagnostics Solutions, Inc).

For the present study we used 46 subjects selected from those with ≥1µg RNA available which showed no / minimal evidence of degradation on visual inspection of a Bioanalyser (Agilent) trace (RNA integrity number calculations are not valid for dual species RNA analysis). To reduce potential confounding we aimed to frequency match subjects between SM and UM groups as closely as possible by age and gender, and if there remained a choice of samples available we selected those with the most complete additional clinical and laboratory data. For UM samples we aimed to include an equal number with parasitemia above and below 5% (to maximise the chance of obtaining parasite reads in some of the subjects). For SM samples we aimed to include subjects with each of the common SM phenotypes seen in this part of the Gambia in approximately equal numbers, although final numbers were determined by availability and quality of RNA. Detailed information about the study subjects is shown in Supplementary Table 1 and Supplementary Dataset 1. Characteristics were compared between subject groups using one-way ANOVA for continuous data and Fisher's exact test for categorical data.

RNA sequencing

Total RNA was extracted using the PAXgene Blood RNA kit (BD). Libraries were prepared from 1µg of total RNA using the ScriptSeq v2 RNA-seq library preparation kit (Illumina) with additional steps to remove ribosmal RNA (rRNA) and globin messenger RNA (mRNA) using the Globin-Zero Gold kit (Epicentre). Strand-specific libraries were sequenced using the 2x100 bp protocol with an Illumina HiSeq 2500 instrument. In order to eliminate batch effects, samples were randomized for the order of library preparation. For sequencing, 5-6 samples were run per lane, and each lane contained at least one sample from each disease type, randomly allocated in a block design. Library preparation and sequencing were carried out by Exeter University sequencing service.

Genomes and RNA annotations

Human reference genome (hg38) was obtained from UCSC genome browser (http://genome.ucsc.edu/) and *P. falciparum* reference genome (release 24) was obtained from PlasmoDB (http://plasmodb.org/). Human gene annotation was obtained from GENCODE (release 22) (http://gencodegenes.org/releases/) and *P. falciparum* gene annotation from PlasmoDB (release 24) (http://plasmodb.org).

Read Mapping and quantification

RNA-seq data was mapped to the combined genomic index containing both human and *P. falciparum* genomes using the splice-aware STAR aligner, allowing up to 8 mismatches for each paired-end read (*82*). Reads were extracted from the output BAM file to separate parasite-mapped reads from human-mapped reads. Reads mapping to both genomes were counted for each sample and removed. BAM files were sorted, read groups replaced with a single new read group and all reads assigned to it, and indexed to run RNA-SeQC, a tool for computing quality control metrics for RNA-seq data (*83*). HTSeq-count was used to count the reads mapped to exons with the parameter "-m union" (*84*). Only uniquely mapping reads were counted.

Since our analysis of *P. falciparum* gene expression was reliant on a reference genome, families of highly polymorphic *var*, *stevor*, and *rifin* genes were removed from downstream analyses as these exhibit great sequence diversity between parasites and are likely to be incorrectly characterized (*14*).

626 Additional highly polymorphic regions within the *P. falciparum* genome which might also be
627 incorrectly characterized were identified using schizont stage RNA-seq data from 9 clinical isolates
628 (Duffy et al., manuscript submitted). In total, 139 genes were identified with highly polymorphic
629 regions. A reference GTF file containing *P. falciparum* gene annotations was modified to remove
630 these regions without removing the genes, and the resulting read count data generated using the
631 modified GTF file was used for downstream analysis.
632 **Outlier identification**
633 With the R package edgeR, raw read counts of each data set were normalized using a trimmed mean
634 of M-values (TMM), which takes into account the library size and the RNA composition of the input
635 data (*85*). A multi-dimensional scaling (MDS) plot was used to identify the distances between
636 samples that correspond to leading biological coefficient of variation. Up to the 6th dimension of
637 MDS was plotted to fully observe the variation between samples, with two dimensions visualized at
638 a time in scatter plot format. Three parasite samples were consistently found to be positioned away
639 from other samples in each pair of dimensions, indicating outliers. This was further supported by low
640 correlations observed between either of these outliers with other samples. These three samples
641 were excluded from further parasite gene expression analysis: one sample (HL_478) had very low
642 parasite reads making estimation of gene expression impossible and the other two samples (CH_285
643 and UM_589) were conspicuous outliers on MDS plots, possibly due to imperfect library preparation.
644 **Deconvolution analysis**
645 In order to study the variations in gene expression beyond those attributable to inter-individual
646 variation in the proportion of different cell types or developmental stage of circulating parasites we
647 sought to determine the contribution of each cell type and parasite stage to the RNA we detected
648 from whole blood, and then adjust for the effect of cell-mixture when analysing the association
649 between gene expression and severity. We expected that this would particularly enhance the
650 detection of cellular responses to external stimuli (upstream regulators) which may otherwise be
651 masked by variation in cell proportions.Deconvolution analysis was performed on RNA-seq data

652 using CellCODE (*25*). This uses a multi-step statistical framework to compute the relative differences
653 in cell proportion represented as surrogate proportion variables (SPVs). It requires a reference data
654 set that contains gene expression profiles for each cell type of interest. Five major immune cell
655 populations were selected from Immune Response In Silico (IRIS) (*86*) to constitute the human
656 reference data set: neutrophil, monocyte, CD4+ T-cell, CD8+ T-cell, and B-cell. Fragments Per
657 Kilobase of transcript per Million mapped reads (FPKM) values were calculated from human RNA-seq
658 data and log-transformed to simulate a microarray data set. For the parasite reference data set,
659 RNA-seq data sets were obtained for four specific stages in the parasite asexual and sexual stage (0
660 hour, 24 hour, 48 hour, and gametocyte stage V) (*26, 28*), normalized by relative library sizes of
661 samples (i.e. size factors) using edgeR. An identical normalization method was also applied for the
662 input parasite RNA-seq data. A trial-and-error approach was taken to obtain the optimum SPV values
663 for each cell-type. For human deconvolution, a cutoff value of 1.2 and a maximum number of marker
664 genes of 50 appeared optimal. For parasite deconvolution, a cutoff value of 1.7 and a maximum
665 number of marker genes of 50 appeared optimal.
666 Validation of CellCODE for *P. falciparum* developmental stage deconvolution (for which its use has
667 not previously been reported) was performed by comparison with previously reported "stage-
668 specific" marker genes (*29*) and by assessing performance in synthetic data sets constructed by
669 mixing together in varying proportions randomly selected reads from RNA-seq reference datasets
670 (*26, 28*) of the different parasite developmental stages.
671 The effect of adjustment for SPVs on the clustering of subjects based on global gene expression was
672 assessed using multivariate ANOVA to compare the first 10 principal components of gene expression
673 in SM vs UM subjects, before and after adjustment for SPVs.
674 In order to examine the potential importance of variation in cell numbers as a component of the
675 host response to infection, we examined the intensity of expression per unit volume of blood of
676 several human genes with relatively cell-specific expression. We multiplied normalized expression

(FPKM) of each gene after deconvolution by the absolute cell count to give a composite measure representing expression per mL of blood.

Differential gene expression and regression analyses

Prior to carrying out any downstream analyses genes with very low TMM-normalized read counts (< 5 counts-per-million (cpm) in < 3 samples and undetected in the remainder) were excluded. The generalized linear model (glm) tool in edgeR was employed to perform differential gene expression analysis (DGEA) between disease groups with adjustment for leukocyte and parasite SPVs, and in subsequent analyses additional adjustment for log parasite density and log PfHRP2.

Linear regression analysis was performed in edgeR to identify genes significantly associated with clinical variables of interest. Input gene expression values included adjustment for SPVs. The variables considered were: log PfHRP2, log parasite density, lactate concentration, platelet counts, and hemoglobin concentration. Additional analysis for lactate, platelets and hemoglobin were conducted including adjustment for log parasite density and log PfHRP2. Association with BCS was assessed using ordinal linear regression with the R package polr, followed by chi-squared test to compare log likelihood values of the gene models with those of a model fitted with a constant.

In both DGEA and linear regression analyses, false Discovery Rate (FDR) was computed for each individual analysis using the Benjamini-Hochberg procedure (*87*). Genes with FDR below 0.05 were considered to be differentially expressed.

Gene ontology and KEGG pathway enrichment analysis

Gene ontology (GO) terms for genes were obtained from Bioconductor package “org.Hs.eg.db” for human and “org.Pf.plasmo.db” for parasite. Input gene lists were significantly differentially expressed genes or genes that were significantly associated with laboratory variables. Fisher’s exact test was used to identify significantly over-represented GO terms from these gene lists. The background sets for each species consisted of all expressed genes detected in the data set with the exclusion of those with very low expression as described above. Enrichment analysis for biological

process terms was carried out using the "goana()" function in edgeR. The least redundant GO terms with greatest significance in each analysis were identified for reporting using the tool REVIGO (*88*). Ingenuity Pathway Analysis (Qiagen) was used for prediction of upstream regulators of groups of differentially expressed genes, and to identify functional networks.

Construction of a co-expression network

The weighted gene co-expression network analysis (WGCNA) tool was used to construct a gene co-expression network (*52*). The input data for WGCNA was read counts for each gene feature normalized using TMM method and then adjusted for SPVs using the command "removeBatchEffect()" from the R package edgeR. In order to comprehensively study the relationships between genes, three separate sets of networks were created. For the first two, paired human and parasite expression data were analyzed together as a single set of genes for each subject. We created one network using all samples from SM and UM groups at the same time, and the other with two separate sub-networks keeping SM and UM groups separate. For the third network we used only human genes for the combined SM and UM subjects. Networks were generated following the WGCNA tool guidelines:

1) Hierarchical clustering was performed at a sample level to detect outliers based on the WGCNA tool threshold. If outliers were identified they were removed from the subsequent network generation (HL_171 and UM_492 were removed from combined networks).

2) An appropriate soft-thresholding power (b) was chosen by applying the scale-free topology criterion. This was such that the power value enables the resulting gene network to satisfy the scale-free topology of approximately ($R^2 > 0.80$).

3) Adjacency, which represents the connection strength of two genes in a network, was calculated. Co-expression similarity was calculated by taking the absolute value of the correlation coefficient, multiplying by 0.5 and adding 0.5 to create a signed network, where the presence of strongly negatively correlated gene pairs is downsized.

4) The adjacency matrix was transformed into a topological overlap matrix (TOM) in order to minimize the effects of spurious associations and noise in the network.

5) Hierarchical clustering on TOM dissimilarity was done to create hierarchical clustering tree of genes.

6) The dynamic tree cut method was used to group the genes that are highly correlated with one another into gene modules where minimum module size and the tree height at which genes below the height is grouped together were specified.

7) Module eigengene values for each module were calculated, which represents the overall gene expression profile of a module. Correlation analysis between modules was performed using eigengene values to identify modules with high similarity (R>0.75), which were then merged together.

The resulting network consisted of genes (represented as nodes in the network) and correlations between genes (represented as edges in the network), and highly correlated genes grouped together into modules. To characterize the gene network, several analysis steps were carried out. The most connected genes in each module were identified as the hub genes. Pearson correlation analysis between module eigengene values and clinical variables was performed to identify gene clusters that are highly associated with clinical traits (Spearmann correlation was used for BCS). Gene set enrichment analysis was performed on each module to identify significantly enriched GO terms. This data was summarised using OmicCircos (*89*).

From two separate sub-networks generated from SM and UM groups respectively, the preservation of gene connections across SM and UM groups was determined by assessing an overlap of genes for each module pair (from SM and UM sub-networks respectively), the significance of overlap was measured using the hypergeometric test. For each significantly preserved module pair, the hub genes and the significantly enriched GO terms were compared.

The gene network was exported to Cytoscape (http://www.cytoscape.org/) for visualisation. Only gene pairs with adjacency value of 0.03 or higher were exported to remove genes with low

753 connections from the network visualization. SM and UM sub-networks were exported separately
754 and subsequently combined into a single network using the Cytoscape embedded tool "Merge". By
755 doing so, duplicate genes representing overlap between SM and UM sub-networks were removed,
756 and connections between genes remained intact such that genes that can only be found on SM sub-
757 network and also connected to the genes that can be found on both networks were not connected
758 to the genes that can be found on UM sub-network and also connected to the same overlapping
759 genes.
760 Correlation analysis between human gene co-expression module eigengene values and parasite gene
761 expression values with adjustment for parasite load was performed using the glm package in R, using
762 only human modules with significant difference in module eigengene values between SM and UM
763 ($P<0.05$). Gene set enrichment analysis was then performed on significantly correlated parasite
764 genes (FDR-adjusted $P<0.05$) from correlations with each module to identify significantly enriched
765 GO terms.
766 **Logistic regression for association of module eigengenes with severity**
767 Logistic regression was performed using the glm package in R to identify module eigengene values
768 with univariate association with severity. All modules with significant univariate associations
769 ($P<0.01$) in addition to log PfHRP2 concentration were used in backward selection to identify the
770 best multivariate model in which all terms were significant.
771 **Lactate supplementation of parasite cultures**
772 *P. falciparum 3D7* strain was grown in human A+ erythrocytes at 1-5% haematocrit in RPMI-1640
773 (without L-glutamine, with HEPES) (Sigma) supplemented with 5 g/liter Albumax II (Invitrogen), 147
774 M hypoxanthine, 2 mM L-glutamine,10 mM D-glucose in 5% CO2 and low oxygen (1% or 5%) at 37
775 °C. Parasites were synchronised at early ring stage through a combination of Percoll & sorbitol
776 treatments and adjusted to 5% parasitemia immediately prior to supplementation with sodium L-
777 15mM lactate or control medium. After 5h erythrocytes were pelleted, adjusted to 50% haematocrit
778 and transferred to Paxgene tubes (Qiagen). RNA extraction, quality control, RNA sequencing and

779 analysis including deconvolution of parasite developmental stage, adjustment for SPVs and
780 differential expression analysis were performed as described above, with the exception that a
781 2x125bp sequencing protocol was used.
782 **Microarray dataset analysis**
783 Published human microarray gene expression data, Gene Expression Omnibus GSE72058 (*43*), was
784 used to assess the contribution of parasite load to differential gene expression between cerebral
785 malaria phenotypes. Reanalysis was restricted to subjects with parasite density and PfHRP2
786 measurements (kindly provided by Johanna Daily, Albert Einstein College of Medicine): 55
787 retinopathy-positive and 17 retinopathy-negative cerebral malaria cases. Deconvolution and
788 adjustment for cell mixture was performed using CellCODE as described above. DGEA was
789 performed using Limma (for microarray) with and without adjustment for log parasite density or log
790 PfHRP2.
791 **Whole blood stimulation assay**
792 Schizonts were enriched from tightly synchronized *P. falciparum* 3D7 parasites by Percoll gradient
793 centrifugation. The schizont pellet (~80% parasitemia) was re-suspended in RPMI to 50% hematocrit
794 and subjected to three rapid freeze-thaw cycles to generate lysate. Whole blood was collected from
795 8 healthy adult donors into sodium heparin tubes, diluted 1:1 in RPMI, and incubated overnight with
796 or without schizont lysate (1:10). Supernatant was collected for further analysis.
797 **Protein measurements**
798 Proteins were measured in stored plasma samples collected on the day of clinical presentation or
799 culture supernatants using ELISAs according to the manufacturers' instructions: MMP8, human
800 neutrophil elastase, GCSF (all Abcam), and defensin A3 (Cloud Clone). Spearman correlation was
801 used to assess correlations between plasma proteins concentrations, gene expression and module
802 eigengene values.
803

Supplementary Material (Separate Supplementary file with captions for those that can be embedded in a word document)

Supplementary Table 1. Human genes differentially expressed between severe malaria phenotypes and uncomplicated malaria in unadjusted and parasite load-adjusted analyses.

Supplementary Table 2. Gene ontology terms associated with human differentially expressed or significantly correlated genes in unadjusted and parasite load-adjusted analyses.

Supplementary Table 3. Predicted upstream regulators associated with human differentially expressed or significantly correlated genes in unadjusted and parasite load-adjusted analyses.

Supplementary Table 4. Human genes significantly correlated with parasite load and pathophysiological variables in unadjusted and parasite load-adjusted analyses.

Supplementary Table 5. *P. falciparum* genes differentially expressed between severe malaria phenotypes and uncomplicated malaria in unadjusted and parasite load-adjusted analyses.

Supplementary Table 6. Gene ontology terms associated with parasite differentially expressed or significantly correlated genes in unadjusted and parasite load-adjusted analyses.

Supplementary Table 7. *P.falciparum* genes significantly correlated with parasite load and pathophysiological variables in unadjusted and parasite load-adjusted analyses.

Supplementary Table 8. Summary of modules obtained from combined whole genome correlation network.

Supplementary Table 9. Univariate and multivariate associations of module eigengene values and parasite load with severity.

Supplementary Table 10. Summary and overlap of whole genome correlation sub-networks for severe and uncomplicated malaria.

Supplementary Table 11. Summary of modules obtained from human only whole genome correlation network.

Supplementary Fig 1. Estimates of the relative proportions of leukocyte subpopulations in subjects with severe and uncomplicated malaria. (A-E) Surrogate proportion variables, calculated using CellCODE, are compared by severity category for neutrophils (**A**), monocytes (**B**), CD4+ T-lymphocytes (**C**), CD8+ T-lymphocytes (**D**), and B-lymphocytes (**E**) using the Mann-Whitney test (UM, n=21; SM, n=25; bold line, box and whiskers indicate median, interquartile range and 1.5-times interquartile range respectively).

Supplementary Fig 2. Validation of the gene-signature approach to estimate parasite developmental stage proportions. (A-C) Correlation of surrogate proportion variables (SPV), calculated using CellCODE, with read counts for putative "stage-specific" marker genes (see Methods): (**A**) 0hr SPV vs. early asexual stage marker gene PFE0065w; (**B**) 48hr SPV vs. late asexual stage marker gene PF10_0020; (**C**) Gametocyte V SPV vs. developing gametocyte marker gene PF14_0367 (R for Pearson correlation, n=43). (**D-G**) Correlation of SPVs with actual proportion of reads derived from each parasite developmental stage in synthetic mixtures of varying proportions of stage-specific RNA-seq reads from early ring-stage (0 hour, **D**), trophozoite (24 hour, **E**), late schizont (48 hour, **F**) and mature gametocyte (Stage V, **G**). R for Pearson correlation, n=43.

Supplementary Fig 3. Differential gene expression between severe malaria phenotypes and uncomplicated malaria. Volcano plots showing extent and significance of up- or down- regulation of human (left hand column) or *P. falciparum* (right hand column) gene expression in comparisons between specific phenotypes of SM vs UM (red and green, P <0.05 after Benjamini-Hochberg adjustment for false discovery rate (FDR); orange and green, absolute log_2-fold change (FC) in expression > 1).

Supplementary Fig 4. Top functional network for the small LYSMD3 module. Functional networks were identified in Ingenuity Pathway Analysis software and the top scoring network is portrayed in radial layout which places the most interconnected gene at the centre. Genes within the module are shaded.

Supplementary Fig 5. **Association between gene expression and plasma protein concentrations.** (A-B) Correlation between plasma protein concentrations and corresponding gene expression of *DEFA3* (A) and *ELANE* (B) (n=34; SM, n=14; UM, n= 20). (C-E) Correlation between plasma MMP8 and *MMP8* expression in all subjects (C, n=35) and separately in those with SM (D, n=15) or UM (E, n=20). (F) Correlation of plasma GCSF (with eigengene values for the MMP8 module derived from the combined network analysis of subjects with SM and UM (analysis restricted to subjects with detectable plasma GCSF, n=19; SM, n=8; UM, n=11). (A-F) P and rho for Spearman correlation. (G) MMP8 release from healthy human adult donor blood (n=8) before and after stimulation with *P. falciparum* 3D7 schizont lysate, P for Wilcoxon matched pairs test.

Supplementary Dataset 1. Subject-level clinical and laboratory data

880 **(After references) Acknowledgements:** We are grateful to the study subjects, staff at MRCG,
881 Jammeh Foundation for Peace Hospital, and Brikama Health Centre; Konrad Paszkiewicz and staff at
882 Exeter Sequencing Service at the University of Exeter; Johanna Daily for providing subject level data
883 for GSE72058. **Funding:** This work was funded by the UK Medical Research Council (MRC) and the UK
884 Department for International Development (DFID) under the MRC/DFID Concordat agreement and is
885 also part of the EDCTP2 programme supported by the European Union (MR/L006529/1), MRC core
886 funding of the MRC Gambia Unit (MRCG), and Wellcome Trust (098051) ; Exeter Sequencing Service
887 is supported by Medical Research Council Clinical Infrastructure award (MR/M008924/1), Wellcome
888 Trust Institutional Strategic Support Fund (WT097835MF), Wellcome Trust Multi User Equipment
889 Award (WT101650MA) and BBSRC LOLA award (BB/K003240/1). **Author contributions:** AJC, ML and
890 DJC conceived the study; HJL, TDO and AG performed formal analysis; MW and AJC performed
891 investigations on clinical samples; MW, DJC, DN and LBS provided samples; DJC, TDO and LBS
892 provided methodology; AJC and HJL wrote the original draft; all authors contributed to review and
893 editing of the manuscript; LJC, DJC, ML, TDO and AJC provided supervision; AJC obtained funding.
894 **Competing interests:** The authors declare no competing financial interests. **Data availability:** RNA-
895 seq data have been deposited in the ArrayExpress database at EMBL-EBI
896 (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-6413. Source data for Table 1 are
897 provided with the paper as Supplementary Dataset 1. Correspondence should be addressed to Dr
898 Aubrey Cunnington, Clinical Senior Lecturer, Section of Paediatrics, Department of Medicine,
899 Imperial College London (St Mary's Campus) 231, Medical School Building, Norfolk Place, London W2
900 1PG. a.cunnington@imperial.ac.uk
901

902 **Table 1. Characteristics of study subjects (n=46)**

	CM (n=5)	CH (n=12)	HL (n=8)	UM (n=21)	*P (F value)*
Age (years)	4.3 (4.2-4.8)	4.9 (3.6-5.7)	5.0 (3.8-8.3)	6.0 (4.0-9.0)	0.24 (1.47)
Male (%)	3 (60%)	5 (42%)	7 (88%)	13 (62%)	0.24
Parasitemia (%)	8.3 (5.3-9.0)[4]	12.6 (9.4-19.0)	9.6 (1.8-12.2)	5.1 (3.8-7.0)	0.02 (3.67)
Parasites (x10^5 /uL)	2.3 (1.7-3.1)[3]	3.5 (2.7-8.4)[11]	2.8 (0.7-5.0)	2.3 (1.6-3.2)	0.02 (3.55)
Clones	2 (1.5-2.5)[4]	2 (1-2)[9]	1 (1-2)[5]	2 (1-2)[15]	0.68 (0.51)
PfHRP2 (ng/mL)	202 (93-528)[4]	763 (374-1750)	470 (164-2214)	163 (128-227)	0.001 (6.20)
Duration of illness (days)	2.0 (1.7-3.0)	2.0 (2.0-2.5)	2.0 (2.0-3.5)	2.7 (2.0-3.0)	0.57 (0.68)
Hb (g/dL)	9.7 (7.4-10.4)	9.3 (7.8-11.5)[11]	9.1 (7.4-11.0)	10.8 (9.9-12.1)	0.10 (2.24)
WBC (x10^9/L)	9.8 (8.2-12.9)[4]	8.8 (6.4-9.4)[11]	15.3 (7.9-16.8)[7]	9.5 (7.7-11.8)	0.30 (1.27)
Platelets (x10^9/L)	41 (40-82)[4]	36 (23-65)[11]	59 (33-132)	122 (96-132)	0.04 (3.15)
Lymphocytes (x10^9/L)	2.7 (2.1-3.6)[4]	2.9 (2.4-3.6)[11]	3.1 (1.8-5.2)[7]	2.4 (1.4-3.1)[20]	0.53 (0.75)
Lymphocyte (%)	29.8 (20.6-37.3)[4]	37.8 (29.9-49.9)[11]	22.3 (14.7-37.3)[7]	23.9 (16.0-33.5)[20]	0.05 (2.78)
Neutrophils (x10^9/L)	6.4 (4.0-8.7)[4]	4.0 (2.9-4.3)[10]	6.5 (5.8-10.4)[7]	7.0 (5.3-7.7)[20]	0.17 (1.79)
Neutrophil (%)	55.1 (49.0-69.6)[4]	48.3 (39.6-56.2)[10]	61.5 (55.6-74.9)[7]	68.0 (59.9-79.6)[20]	0.01 (4.25)
Monocytes (x10^9/L)	0.6 (0.6-0.7)[4]	0.6 (0.5-0.9)[10]	0.8 (0.6-1.3)[7]	0.7 (0.4-0.9)[20]	0.53 (0.75)
Monocyte (%)	7.1 (6.0-7.7)[4]	7.8 (6.8-8.6)[10]	6.6 (5.1-7.8)[7]	6.7 (4.8-7.3)[20]	0.38 (1.05)

903 CM, cerebral malaria; CH, cerebral malaria plus hyperlactatemia; HL, hyperlactatemia (CM, CH, and
904 HL are all subgroups of severe malaria, SM); UM, uncomplicated malaria; PfHRP2, *P. falciparum*
905 histidine rich protein 2; Hb, hemoglobin concentration; WBC, white blood cell count. Data are
906 median (IQR), superscripts indicate the number of subjects with data for each variable if less than

907 the total; *P (F)* for ANOVA comparing all groups (degrees of freedom =3) except for sex where *P* is for
908 Fisher's exact test. Leukocyte population numbers and proportions measured by clinical hematology
909 analyser.
910

Table 2. Numbers of differentially expressed genes before and after adjustment for parasite load

	Human genes				Parasite genes			
	n*	Unadjusted	Log parasite density	Log PfHRP2	n*	Unadjusted	Log parasite density	Log PfHRP2
SM vs UM	43	907	914 (101%)	13 (1.4%)	41	516	562 (109%)	329 (64%)
BCS	43	738	491 (67%)	12 (1.6%)	41	445	340 (76%)	148 (33%)
Lactate	40	1012	526 (52%)	51 (5.0%)	38	100	109 (109%)	47 (47%)
Platelets	43	178	66 (37%)	46 (25%)	41	1	1 (100%)	1 (100%)
Hemoglobin	43	0	0	0	41	6	5 (83%)	4 (67%)

*Only subjects with complete data for every parameter are included. Number of genes associated with severity category or laboratory marker of severity before and after adjustment for parasite load (% of number in unadjusted analysis where applicable). BCS, Blantyre coma scale.

Table 3. Parasite genes correlated with human gene co-expression modules after adjustment for parasite load

Human Module		Correlated parasite genes after adjustment for parasite load				
Hub gene	**Top GO term**	**n (+/-)**	**Top genes**	**FDR *P***	**Top GO terms**	***P***
Receptor Transporter Protein 4 (RTP4)	GO:0034340 response to type I interferon	2 (2/0)	*PF3D7_0827500* (apicoplast ribosomal protein L21 precursor)	0.014	GO:0006412 translation	0.018
			PF3D7_0111800 (eukaryotic translation initiation factor 4E)	0.035		
Trinucleotide Repeat Containing 6B (TNRC6B)	GO:0016569 chromatin modification	97 (29/68)	*PF3D7_1119200* (unknown function)	0.00012	GO:0008380 RNA splicing	4.15×10^{-6}
			PF3D7_1309100 (60S ribosomal protein L24)	0.00027	GO:0006396 RNA processing	3.88×10^{-5}
			PF3D7_0825500 (protein KRI1)	0.00036		
Heat shock protein family H (Hsp110) member 1 (HSPH1)	GO:0006457 protein folding	21 (17/4)	*PF3D7_0933100* (unknown function)	0.0071	GO:0000338 protein deneddylation	0.0038
			PF3D7_1118400 (haloacid dehalogenase-like hydrolase)	0.0079		
			PF3D7_0521800 (AFG1-like ATPase)	0.01		
Matrix metallopeptid-ase 8 (MMP8)	GO:0009617 response to bacterium	18 (16/2)	*PF3D7_1356200* (mitochondrial import inner membrane translocase subunit TIM23)	0.0045	GO:0019219 regulation of nucleobase-containing	0.0065

			PF3D7_1119100 (tRNA m(1)G methyltransferase)	0.007 1	compound metabolic process	
			PF3D7_0823100 (RWD domain-containing protein)	0.007 3		
Peptidylprolyl isomerase B (PPIB)	GO:0034976 response to endoplasmic reticulum stress	24 (24/ 0)	*PF3D7_1420300* (DNL-type zinc finger protein)	0.002 2	GO:0006364 rRNA processing	0.0058
			PF3D7_0821200 (unknown function)	0.005 7		
			PF3D7_1119200 (unknown function)	0.005 7		
Ribosomal protein L24 (RPL24)	GO:0006614 SRP-dependent co-translational protein targeting to membrane	93 (78/ 15)	*PF3D7_0530600* (XAP-5 DNA binding protein)	5.3 $x10^{-6}$	GO:0006396 RNA processing	6.92 $x10^{-4}$
			PF3D7_1309100 (60S ribosomal protein L24)	5.3 $x10^{-6}$	GO:0008380 RNA splicing	1.10 $x10^{-3}$
			PF3D7_0821200 (unknown function)	3.3 $x10^{-5}$		

+/- indicates number of parasite genes positively / negatively correlated with each human module eigengene value. All "top genes" in the table are positively correlated with the module eigengene value.

References

1. A. J. Westermann, L. Barquist, J. Vogel, Resolving host-pathogen interactions by dual RNA-seq. *PLoS Pathog* **13**, e1006033 (2017).
2. A. J. Westermann, S. A. Gorski, J. Vogel, Dual RNA-seq of pathogen and host. *Nat Rev Microbiol* **10**, 618-630 (2012).
3. M. A. Phillips, J. N. Burrows, C. Manyando, R. H. van Huijsduijnen, W. C. Van Voorhis, T. N. C. Wells, Malaria. *Nat Rev Dis Primers* **3**, 17050 (2017).
4. A. J. Cunnington, M. Walther, E. M. Riley, Piecing together the puzzle of severe malaria. *Sci Transl Med* **5**, 211ps218 (2013).
5. S. C. Wassmer, G. E. Grau, Severe malaria: what's new on the pathogenesis front? *Int J Parasitol* **47**, 145-152 (2017).
6. K. Marsh, D. Forster, C. Waruiru, I. Mwangi, M. Winstanley, V. Marsh, C. Newton, P. Winstanley, P. Warn, N. Peshu, et al., Indicators of life-threatening malaria in African children. *N Engl J Med* **332**, 1399-1404 (1995).
7. J. M. O'Sullivan, R. J. Preston, N. O'Regan, J. S. O'Donnell, Emerging roles for hemostatic dysfunction in malaria pathogenesis. *Blood* **127**, 2281-2288 (2016).
8. A. M. Dondorp, V. Desakorn, W. Pongtavornpinyo, D. Sahassananda, K. Silamut, K. Chotivanich, P. N. Newton, P. Pitisuttithum, A. M. Smithyman, N. J. White, N. P. Day, Estimation of the total parasite biomass in acute falciparum malaria from plasma PfHRP2. *PLoS Med* **2**, e204 (2005).
9. I. C. Hendriksen, J. Mwanga-Amumpaire, L. von Seidlein, G. Mtove, L. J. White, R. Olaosebikan, S. J. Lee, A. K. Tshefu, C. Woodrow, B. Amos, C. Karema, S. Saiwaew, K. Maitland, E. Gomes, W. Pan-Ngum, S. Gesase, K. Silamut, H. Reyburn, S. Joseph, K. Chotivanich, C. I. Fanello, N. P. Day, N. J. White, A. M. Dondorp, Diagnosing severe falciparum malaria in parasitaemic African children: a prospective evaluation of plasma PfHRP2 measurement. *PLoS Med* **9**, e1001297 (2012).
10. I. C. Hendriksen, L. J. White, J. Veenemans, G. Mtove, C. Woodrow, B. Amos, S. Saiwaew, S. Gesase, B. Nadjm, K. Silamut, S. Joseph, K. Chotivanich, N. P. Day, L. von Seidlein, H. Verhoef, H. Reyburn, N. J. White, A. M. Dondorp, Defining falciparum-malaria-attributable severe febrile illness in moderate-to-high transmission settings on the basis of plasma PfHRP2 concentration. *J Infect Dis* **207**, 351-361 (2013).
11. A. J. Cunnington, M. T. Bretscher, S. I. Nogaro, E. M. Riley, M. Walther, Comparison of parasite sequestration in uncomplicated and severe childhood Plasmodium falciparum malaria. *J Infect* **67**, 220-230 (2013).
12. K. Haldar, S. C. Murphy, D. A. Milner, T. E. Taylor, Malaria: mechanisms of erythrocytic infection and pathological correlates of severe disease. *Annu Rev Pathol* **2**, 217-249 (2007).
13. A. J. Cunnington, E. M. Riley, M. Walther, Stuck in a rut? Reconsidering the role of parasite sequestration in severe malaria syndromes. *Trends Parasitol* **29**, 585-592 (2013).
14. M. Wahlgren, S. Goel, R. R. Akhouri, Variant surface antigens of Plasmodium falciparum and their roles in severe malaria. *Nat Rev Microbiol* **15**, 479-491 (2017).
15. J. Dunst, F. Kamena, K. Matuschewski, Cytokines and Chemokines in Cerebral Malaria Pathogenesis. *Front Cell Infect Microbiol* **7**, 324 (2017).
16. S. W. Howland, C. Claser, C. M. Poh, S. Y. Gun, L. Renia, Pathogenic CD8+ T cells in experimental cerebral malaria. *Semin Immunopathol* **37**, 221-231 (2015).
17. S. J. Higgins, K. C. Kain, W. C. Liles, Immunopathogenesis of falciparum malaria: implications for adjunctive therapy in the management of severe and cerebral malaria. *Expert Rev Anti Infect Ther* **9**, 803-819 (2011).
18. R. T. Gazzinelli, P. Kalantari, K. A. Fitzgerald, D. T. Golenbock, Innate sensing of malaria parasites. *Nat Rev Immunol* **14**, 744-757 (2014).

19. T. van der Poll, F. L. van de Veerdonk, B. P. Scicluna, M. G. Netea, The immunopathology of sepsis and potential therapeutic targets. *Nat Rev Immunol* **17**, 407-420 (2017).
20. J. B. Prescott, A. Marzi, D. Safronetz, S. J. Robertson, H. Feldmann, S. M. Best, Immunobiology of Ebola and Lassa virus infections. *Nat Rev Immunol* **17**, 195-207 (2017).
21. P. J. M. Openshaw, C. Chiu, F. J. Culley, C. Johansson, Protective and Harmful Immunity to RSV Infection. *Annu Rev Immunol* **35**, 501-532 (2017).
22. A. Scholzen, R. W. Sauerwein, Immune activation and induction of memory: lessons learned from controlled human malaria infection with Plasmodium falciparum. *Parasitology* **143**, 224-235 (2016).
23. J. Yamagishi, A. Natori, M. E. Tolba, A. E. Mongan, C. Sugimoto, T. Katayama, S. Kawashima, W. Makalowski, R. Maeda, Y. Eshita, J. Tuda, Y. Suzuki, Interactive transcriptome analysis of malaria patients and infecting Plasmodium falciparum. *Genome Res* **24**, 1433-1444 (2014).
24. S. S. Shen-Orr, R. Gaujoux, Computational deconvolution: extracting cell type-specific information from heterogeneous samples. *Curr Opin Immunol* **25**, 571-578 (2013).
25. M. Chikina, E. Zaslavsky, S. C. Sealfon, CellCODE: a robust latent variable approach to differential expression analysis for heterogeneous cell populations. *Bioinformatics* **31**, 1584-1591 (2015).
26. T. D. Otto, D. Wilinski, S. Assefa, T. M. Keane, L. R. Sarry, U. Bohme, J. Lemieux, B. Barrell, A. Pain, M. Berriman, C. Newbold, M. Llinas, New insights into the blood-stage transcriptome of Plasmodium falciparum using RNA-Seq. *Mol Microbiol* **76**, 12-24 (2010).
27. W. A. Hoeijmakers, R. Bartfai, H. G. Stunnenberg, Transcriptome analysis using RNA-Seq. *Methods Mol Biol* **923**, 221-239 (2013).
28. M. J. Lopez-Barragan, J. Lemieux, M. Quinones, K. C. Williamson, A. Molina-Cruz, K. Cui, C. Barillas-Mury, K. Zhao, X. Z. Su, Directional gene expression and antisense transcripts in sexual and asexual stages of Plasmodium falciparum. *BMC Genomics* **12**, 587 (2011).
29. R. Joice, V. Narasimhan, J. Montgomery, A. B. Sidhu, K. Oh, E. Meyer, W. Pierre-Louis, K. Seydel, D. Milner, K. Williamson, R. Wiegand, D. Ndiaye, J. Daily, D. Wirth, T. Taylor, C. Huttenhower, M. Marti, Inferring developmental stage composition from gene expression in human malaria. *PLoS Comput Biol* **9**, e1003392 (2013).
30. M. Nacher, P. Singhasivanon, U. Silachamroon, S. Treeprasertsuk, T. Tosukhowong, S. Vannaphan, F. Gay, D. Mazier, S. Looareesuwan, Decreased hemoglobin concentrations, hyperparasitemia, and severe malaria are associated with increased Plasmodium falciparum gametocyte carriage. *J Parasitol* **88**, 97-101 (2002).
31. J. B. Cowland, N. Borregaard, Granulopoiesis and granules of human neutrophils. *Immunol Rev* **273**, 11-28 (2016).
32. A. G. Maier, M. Rug, M. T. O'Neill, M. Brown, S. Chakravorty, T. Szestak, J. Chesson, Y. Wu, K. Hughes, R. L. Coppel, C. Newbold, J. G. Beeson, A. Craig, B. S. Crabb, A. F. Cowman, Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes. *Cell* **134**, 48-61 (2008).
33. T. F. de Koning-Ward, M. W. Dixon, L. Tilley, P. R. Gilson, Plasmodium species: master renovators of their host cells. *Nat Rev Microbiol* **14**, 494-507 (2016).
34. A. Sinha, K. R. Hughes, K. K. Modrzynska, T. D. Otto, C. Pfander, N. J. Dickens, A. A. Religa, E. Bushell, A. L. Graham, R. Cameron, B. F. C. Kafsack, A. E. Williams, M. Llinas, M. Berriman, O. Billker, A. P. Waters, A cascade of DNA-binding proteins for sexual commitment and development in Plasmodium. *Nature* **507**, 253-257 (2014).
35. K. Modrzynska, C. Pfander, L. Chappell, L. Yu, C. Suarez, K. Dundas, A. R. Gomes, D. Goulding, J. C. Rayner, J. Choudhary, O. Billker, A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the Plasmodium Life Cycle. *Cell Host Microbe* **21**, 11-22 (2017).

36. B. Balu, N. Singh, S. P. Maher, J. H. Adams, A genetic screen for attenuated growth identifies genes crucial for intraerythrocytic development of Plasmodium falciparum. *PLoS One* **5**, e13282 (2010).
37. F. K. Glenister, R. L. Coppel, A. F. Cowman, N. Mohandas, B. M. Cooke, Contribution of parasite proteins to altered mechanical properties of malaria-infected red blood cells. *Blood* **99**, 1060-1063 (2002).
38. B. M. Cooke, F. K. Glenister, N. Mohandas, R. L. Coppel, Assignment of functional roles to parasite proteins in malaria-infected red blood cells by competitive flow-based adhesion assay. *Br J Haematol* **117**, 203-211 (2002).
39. H. Ishioka, A. Ghose, P. Charunwatthana, R. Maude, K. Plewes, H. Kingston, B. Intharabut, C. Woodrow, K. Chotivanich, A. A. Sayeed, M. U. Hasan, N. P. Day, A. Faiz, N. J. White, A. Hossain, A. M. Dondorp, Sequestration and Red Cell Deformability as Determinants of Hyperlactatemia in Falciparum Malaria. *J Infect Dis* **213**, 788-793 (2016).
40. L. Mancio-Silva, K. Slavic, M. T. Grilo Ruivo, A. R. Grosso, K. K. Modrzynska, I. M. Vera, J. Sales-Dias, A. R. Gomes, C. R. MacPherson, P. Crozet, M. Adamo, E. Baena-Gonzalez, R. Tewari, M. Llinas, O. Billker, M. M. Mota, Nutrient sensing modulates malaria parasite virulence. *Nature* **547**, 213-216 (2017).
41. M. J. Griffiths, M. J. Shafi, S. J. Popper, C. A. Hemingway, M. M. Kortok, A. Wathen, K. A. Rockett, R. Mott, M. Levin, C. R. Newton, K. Marsh, D. A. Relman, D. P. Kwiatkowski, Genomewide analysis of the host response to malaria in Kenyan children. *J Infect Dis* **191**, 1599-1611 (2005).
42. Y. Idaghdour, J. Quinlan, J. P. Goulet, J. Berghout, E. Gbeha, V. Bruat, T. de Malliard, J. C. Grenier, S. Gomez, P. Gros, M. C. Rahimy, A. Sanni, P. Awadalla, Evidence for additive and interaction effects of host genotype and infection in malaria. *Proc Natl Acad Sci U S A* **109**, 16786-16793 (2012).
43. C. M. Feintuch, A. Saidi, K. Seydel, G. Chen, A. Goldman-Yassen, N. K. Mita-Mendoza, R. S. Kim, P. S. Frenette, T. Taylor, J. P. Daily, Activated Neutrophils Are Associated with Pediatric Cerebral Malaria Vasculopathy in Malawian Children. *MBio* **7**, e01300-01315 (2016).
44. R. E. Vandenbroucke, E. Dejonckheere, P. Van Lint, D. Demeestere, E. Van Wonterghem, I. Vanlaere, L. Puimege, F. Van Hauwermeiren, R. De Rycke, C. Mc Guire, C. Campestre, C. Lopez-Otin, P. Matthys, G. Leclercq, C. Libert, Matrix metalloprotease 8-dependent extracellular matrix cleavage at the blood-CSF barrier contributes to lethality during systemic inflammatory diseases. *J Neurosci* **32**, 9805-9816 (2012).
45. A. Schubert-Unkmeir, C. Konrad, H. Slanina, F. Czapek, S. Hebling, M. Frosch, Neisseria meningitidis induces brain microvascular endothelial cell detachment from the matrix and cleavage of occludin: a role for MMP-8. *PLoS Pathog* **6**, e1000874 (2010).
46. M. Sasai, H. Oshiumi, M. Matsumoto, N. Inoue, F. Fujita, M. Nakanishi, T. Seya, Cutting Edge: NF-kappaB-activating kinase-associated protein 1 participates in TLR3/Toll-IL-1 homology domain-containing adapter molecule-1-mediated IFN regulatory factor 3 activation. *J Immunol* **174**, 27-30 (2005).
47. J. P. Mooney, S. C. Wassmer, J. C. Hafalla, Type I Interferon in Malaria: A Balancing Act. *Trends Parasitol* **33**, 257-260 (2017).
48. R. Tahar, C. Albergaria, N. Zeghidour, V. F. Ngane, L. K. Basco, C. Roussilhon, Plasma levels of eight different mediators and their potential as biomarkers of various clinical malaria conditions in African children. *Malar J* **15**, 337 (2016).
49. P. Chimma, C. Roussilhon, P. Sratongno, R. Ruangveerayuth, K. Pattanapanyasat, J. L. Perignon, D. J. Roberts, P. Druilhe, A distinct peripheral blood monocyte phenotype is associated with parasite inhibitory activity in acute uncomplicated Plasmodium falciparum malaria. *PLoS Pathog* **5**, e1000631 (2009).

50. P. Hamon, P. L. Loyher, C. Baudesson de Chanville, F. Licata, C. Combadiere, A. Boissonnas, CX3CR1-dependent endothelial margination modulates Ly6C(high) monocyte systemic deployment upon inflammation in mice. *Blood* **129**, 1296-1307 (2017).
51. T. F. Fuller, A. Ghazalpour, J. E. Aten, T. A. Drake, A. J. Lusis, S. Horvath, Weighted gene coexpression network analysis strategies applied to mouse weight. *Mamm Genome* **18**, 463-472 (2007).
52. P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
53. A. Nakashima, K. Tanimura-Ito, N. Oshiro, S. Eguchi, T. Miyamoto, A. Momonami, S. Kamada, K. Yonezawa, U. Kikkawa, A positive role of mammalian Tip41-like protein, TIPRL, in the amino-acid dependent mTORC1-signaling pathway through interaction with PP2A. *FEBS Lett* **587**, 2924-2929 (2013).
54. V. Marcil, E. Seidman, D. Sinnett, F. Boudreau, F. P. Gendron, J. F. Beaulieu, D. Menard, L. P. Precourt, D. Amre, E. Levy, Modification in oxidative stress, inflammation, and lipoprotein assembly in response to hepatocyte nuclear factor 4alpha knockdown in intestinal epithelial cells. *J Biol Chem* **285**, 40448-40460 (2010).
55. A. J. Cunnington, The importance of pathogen load. *PLoS Pathog* **11**, e1004563 (2015).
56. M. J. Mackinnon, J. Li, S. Mok, M. M. Kortok, K. Marsh, P. R. Preiser, Z. Bozdech, Comparative transcriptional and genomic analysis of Plasmodium falciparum field isolates. *PLoS Pathog* **5**, e1000644 (2009).
57. G. D. Norata, G. Caligiuri, T. Chavakis, G. Matarese, M. G. Netea, A. Nicoletti, L. A. O'Neill, F. M. Marelli-Berg, The Cellular and Molecular Basis of Translational Immunometabolism. *Immunity* **43**, 421-434 (2015).
58. L. A. O'Neill, E. J. Pearce, Immunometabolism governs dendritic cell and macrophage function. *J Exp Med* **213**, 15-23 (2016).
59. N. H. Hunt, H. J. Ball, A. M. Hansen, L. T. Khaw, J. Guo, S. Bakmiwewa, A. J. Mitchell, V. Combes, G. E. Grau, Cerebral malaria: gamma-interferon redux. *Front Cell Infect Microbiol* **4**, 113 (2014).
60. T. A. Fuchs, A. A. Bhandari, D. D. Wagner, Histones induce rapid and profound thrombocytopenia in mice. *Blood* **118**, 3708-3714 (2011).
61. Y. Alhamdi, S. T. Abrams, S. Lane, G. Wang, C. H. Toh, Histone-Associated Thrombocytopenia in Patients Who Are Critically Ill. *JAMA* **315**, 817-819 (2016).
62. B. Amulic, C. Cazalet, G. L. Hayes, K. D. Metzler, A. Zychlinsky, Neutrophil function: from mechanisms to disease. *Annu Rev Immunol* **30**, 459-489 (2012).
63. N. Berens-Riha, I. Kroidl, M. Schunk, M. Alberer, M. Beissner, M. Pritsch, A. Kroidl, G. Froschl, I. Hanus, G. Bretzel, F. von Sonnenburg, H. D. Nothdurft, T. Loscher, K. H. Herbinger, Evidence for significant influence of host immunity on changes in differential blood count during malaria. *Malar J* **13**, 155 (2014).
64. V. S. Baker, G. E. Imade, N. B. Molta, P. Tawde, S. D. Pam, M. O. Obadofin, S. A. Sagay, D. Z. Egah, D. Iya, B. B. Afolabi, M. Baker, K. Ford, R. Ford, K. H. Roux, T. C. Keller, 3rd, Cytokine-associated neutrophil extracellular traps and antinuclear antibodies in Plasmodium falciparum infected children under six years of age. *Malar J* **7**, 41 (2008).
65. A. Dietmann, R. Helbok, P. Lackner, S. Issifou, B. Lell, P. B. Matsiegui, M. Reindl, E. Schmutzhard, P. G. Kremsner, Matrix metalloproteinases and their tissue inhibitors (TIMPs) in Plasmodium falciparum malaria: serum levels of TIMP-1 are associated with disease severity. *J Infect Dis* **197**, 1614-1620 (2008).
66. S. K. Jorch, P. Kubes, An emerging role for neutrophil extracellular traps in noninfectious disease. *Nat Med* **23**, 279-287 (2017).
67. M. S. Oakley, V. Anantharaman, T. M. Venancio, H. Zheng, B. Mahajan, V. Majam, T. F. McCutchan, T. G. Myers, L. Aravind, S. Kumar, Molecular correlates of experimental cerebral malaria detectable in whole blood. *Infect Immun* **79**, 1244-1253 (2011).

68. A. Porcherie, C. Mathieu, R. Peronet, E. Schneider, J. Claver, P. H. Commere, H. Kiefer-Biasizzo, H. Karasuyama, G. Milon, M. Dy, J. P. Kinet, J. Louis, U. Blank, S. Mecheri, Critical role of the neutrophil-associated high-affinity receptor for IgE in the pathogenesis of experimental cerebral malaria. *J Exp Med* **208**, 2225-2236 (2011).
69. M. Krupka, K. Seydel, C. M. Feintuch, K. Yee, R. Kim, C. Y. Lin, R. B. Calder, C. Petersen, T. Taylor, J. Daily, Mild Plasmodium falciparum malaria following an episode of severe malaria is associated with induction of the interferon pathway in Malawian children. *Infect Immun* **80**, 1150-1155 (2012).
70. A. Haque, S. E. Best, A. Ammerdorffer, L. Desbarrieres, M. M. de Oca, F. H. Amante, F. de Labastida Rivera, P. Hertzog, G. M. Boyle, G. R. Hill, C. R. Engwerda, Type I interferons suppress CD4(+) T-cell-dependent parasite control during blood-stage Plasmodium infection. *Eur J Immunol* **41**, 2688-2698 (2011).
71. A. Haque, S. E. Best, M. Montes de Oca, K. R. James, A. Ammerdorffer, C. L. Edwards, F. de Labastida Rivera, F. H. Amante, P. T. Bunn, M. Sheel, I. Sebina, M. Koyama, A. Varelias, P. J. Hertzog, U. Kalinke, S. Y. Gun, L. Renia, C. Ruedl, K. P. MacDonald, G. R. Hill, C. R. Engwerda, Type I IFN signaling in CD8- DCs impairs Th1-dependent malaria immunity. *J Clin Invest* **124**, 2483-2496 (2014).
72. S. Sharma, R. B. DeOliveira, P. Kalantari, P. Parroche, N. Goutagny, Z. Jiang, J. Chan, D. C. Bartholomeu, F. Lauw, J. P. Hall, G. N. Barber, R. T. Gazzinelli, K. A. Fitzgerald, D. T. Golenbock, Innate immune recognition of an AT-rich stem-loop DNA motif in the Plasmodium falciparum genome. *Immunity* **35**, 194-207 (2011).
73. R. A. Zander, J. J. Guthmiller, A. C. Graham, R. L. Pope, B. E. Burke, D. J. Carr, N. S. Butler, Type I Interferons Induce T Regulatory 1 Responses and Restrict Humoral Immunity during Experimental Malaria. *PLoS Pathog* **12**, e1005945 (2016).
74. G. A. Josling, M. Llinas, Sexual development in Plasmodium parasites: knowing when it's time to commit. *Nat Rev Microbiol* **13**, 573-587 (2015).
75. R. J. Arguello, C. Rodriguez Rodrigues, E. Gatti, P. Pierre, Protein synthesis regulation, a pillar of strength for innate immunity? *Curr Opin Immunol* **32**, 28-35 (2015).
76. S. S. Vembar, D. Droll, A. Scherf, Translational regulation in blood stages of the malaria parasite Plasmodium spp.: systems-wide studies pave the way. *Wiley Interdiscip Rev RNA* **7**, 772-792 (2016).
77. W. Wong, X. C. Bai, B. E. Sleebs, T. Triglia, A. Brown, J. K. Thompson, K. E. Jackson, E. Hanssen, D. S. Marapana, I. S. Fernandez, S. A. Ralph, A. F. Cowman, S. H. W. Scheres, J. Baum, Mefloquine targets the Plasmodium falciparum 80S ribosome to inhibit protein synthesis. *Nat Microbiol* **2**, 17031 (2017).
78. J. E. Lemieux, N. Gomez-Escobar, A. Feller, C. Carret, A. Amambua-Ngwa, R. Pinches, F. Day, S. A. Kyes, D. J. Conway, C. C. Holmes, C. I. Newbold, Statistical estimation of cell-cycle progression and lineage commitment in Plasmodium falciparum reveals a homogeneous pattern of transcription in ex vivo culture. *Proc Natl Acad Sci U S A* **106**, 7559-7564 (2009).
79. K. G. Pelle, K. Oh, K. Buchholz, V. Narasimhan, R. Joice, D. A. Milner, N. M. Brancucci, S. Ma, T. S. Voss, K. Ketman, K. B. Seydel, T. E. Taylor, N. S. Barteneva, C. Huttenhower, M. Marti, Transcriptional profiling defines dynamics of parasite tissue sequestration during malaria infection. *Genome Med* **7**, 19 (2015).
80. M. Walther, D. Jeffries, O. C. Finney, M. Njie, A. Ebonyi, S. Deininger, E. Lawrence, A. Ngwa-Amambua, S. Jayasooriya, I. H. Cheeseman, N. Gomez-Escobar, J. Okebe, D. J. Conway, E. M. Riley, Distinct roles for FOXP3 and FOXP3 CD4 T cells in regulating cellular immunity to uncomplicated and severe Plasmodium falciparum malaria. *PLoS Pathog* **5**, e1000364 (2009).
81. M. Walther, A. De Caul, P. Aka, M. Njie, A. Amambua-Ngwa, B. Walther, I. M. Predazzi, A. Cunnington, S. Deininger, E. N. Takem, A. Ebonyi, S. Weis, R. Walton, S. Rowland-Jones, G. Sirugo, S. M. Williams, D. J. Conway, HMOX1 gene promoter alleles and high HO-1 levels are associated with severe malaria in Gambian children. *PLoS Pathog* **8**, e1002579 (2012).

82. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
83. D. S. DeLuca, J. Z. Levin, A. Sivachenko, T. Fennell, M. D. Nazaire, C. Williams, M. Reich, W. Winckler, G. Getz, RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* **28**, 1530-1532 (2012).
84. S. Anders, P. T. Pyl, W. Huber, HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169 (2015).
85. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).
86. A. R. Abbas, D. Baldwin, Y. Ma, W. Ouyang, A. Gurney, F. Martin, S. Fong, M. van Lookeren Campagne, P. Godowski, P. M. Williams, A. C. Chan, H. F. Clark, Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun* **6**, 319-331 (2005).
87. Y. Benjamini, Y. Hochberg, Controlling False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Statistical Methodology* **57**, 289-300 (1995).
88. F. Supek, M. Bosnjak, N. Skunca, T. Smuc, REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, e21800 (2011).
89. Y. Hu, C. Yan, C. H. Hsu, Q. R. Chen, K. Niu, G. A. Komatsoulis, D. Meerzaman, OmicCircos: A Simple-to-Use R Package for the Circular Visualization of Multidimensional Omics Data. *Cancer Inform* **13**, 13-20 (2014).